# WO 00/50639

# Continued

Conventional clinical trial statistics include hypothesis testing and descriptive methods, as elaborated below. Guidance in the selection of appropriate statistical tests for a particular data set is provided in texts such as: Biostatistics: A Foundation for Analysis in the Health Sciences, 7th edition (Wiley Series in Probability and Mathematical Statistics, Applied Probability and statistics) by Wayne W. Daniel, John Wiley & Sons, 1998; Bayesian Methods and Ethics in a Clinical Trial Design (Wiley Series in Probability and Mathematical Statistics. Applied Probability Section) by J. B. Kadane (Editor), John Wiley & Sons, 1996. Examples of specific hypothesis testing and descriptive statistical procedures that may be useful in analyzing clinical trial data are listed below.

A. Hypothesis testing statistical procedures

(1) One-sample procedures (binomial confidence interval, Wilcoxon signed rank test, permutation test with general scores, generation of exact permutational distributions)

(2) Two-sample procedures (*t*-test, Wilcoxon-Mann-Whitney test, Normal score test, Median test, Van der Waerden test, Savage test, Logrank test for censored survival data, Wilcoxon-Gehan test for censored survival data, Cochran-Armitage trend test, permutation test with general scores, generation of exact permutational distributions)

(3) R x C contingency tables (Fisher's exact test, Pearson's chi-squared test, Likelihood ratio test, Kruskal-Wallis test, Jonckheere-Terpstra test, Linear-by linear association test, McNemar's test, marginal homogeneity test for matched pairs)

(4) Stratified 2 x 2 contingency tables (test of homogeneity for odds ratio, test of unity for the common odds ratio, confidence interval for the common odds ratio)

(5) Stratified 2 x C contingency tables (all two-sample procedures listed above with stratification, confidence intervals for the odds ratios and trend, generation of exact permutational distributions)

(6) General linear models (simple regression, multiple regression, analysis of variance –ANOVA-, analysis of covariance, response-surface models, weighted regression, polynomial regression, partial correlation, multiple analysis of variance -MANOVA-, repeated measures analysis of variance).

(7) Analysis of variance and covariance with a nested (hierarchical) structure.

(8) Designs and randomized plans for nested and crossed experiments (completely randomized design for two treatment, split-splot design, hierarchical design, incomplete block design, latin square design)

(9) Nonlinear regression models

(10) Logistic regression for unstratified or stratified data, for binary or ordinal response data, using the logit link function, the normit function or the complementary log-log function.

(11) Probit, logit, ordinal logistic and gompit regression models.

(12) Fitting parametric models to failure time data that may be right-, left-, or interval-censored. Tested distributions can include extreme value, normal and logistic distributions, and, by using a log transformation, exponential, Weibull, lognormal, loglogistic and gamma distributions.

(13) Compute non-parametric estimates of survival distribution with right-censored data and compute rank tests for association of the response variable with other variables.

B. Descriptive statistical methods

- Factor analysis with rotations
- Canonical correlation
- Principal component analysis for quantitative variables.
- Principal component analysis for qualitative data.
- Hierarchical and dynamic clustering methods to create tree structure, dendrogram or phenogram.
- Simple and multiple correspondence analysis using a contingency table as input or raw categorical data.

Specific instructions and computer programs for performing the above calculations can be obtained from companies such as: SAS/STAT Software, SAS Institute Inc., Cary, NC, USA; BMDP Statistical Software, BMDP Statistical Software Inc., Los Angeles, CA, USA; SYSTAT software, SPSS Inc., Chicago, IL, USA; StatXact & LogXact, CYTEL Software Corporation, Cambridge, MA, USA.

C. Statistical Genetic Methods Useful for Analysis of Pharmacogenetic Data

A wide spectrum of mathematical and statistical tools may be useful in the analysis of data produced in pharmacogenetic clinical trials, including methods employed in molecular, population, and quantitative genetics, as well as genetic epidemiology. Methods developed for plant and animal breeding may be useful as well, particularly methods relating to the genetic analysis of quantitative traits.

Analytical methods useful in the analysis of genetic variation among individuals, populations and species of various organisms are described in the following texts: Molecular Evolution, by W- H. Li, Sinauer Associates, Inc., 1997; Principles of Population Genetics, by D. L. Hartl and A. G. Clark, 1996; Genetics and Analysis of Quantitative Traits, By M. Lynch and B. Walsh, Sinauer Associates, Inc., Principles of Quantitative Genetics, by D. S. Falconer and T.F.C. Mackay, Longman, 1996; Genetic Variation and Human Disease, by K. M. Weiss, Cambridge University Press, 1993; Fundamentals of Genetic Epidemiology, by M. J. Khoury, T. H. Beaty, and B. H. Cohen, Oxford University Press, 1993; Handbook of Genetic Linkage, by J. Terwilliger J. Ott, Johns Hopkins University Press, 1994.

The types of statistical analysis performed in different branches of genetics are outlined below as a guide to the relevant literature and publicly available software, some of which is cited.

*Molecular evolutionary genetics*

- Patterns of nucleotide variation among individuals, families/populations and across species and genera,
- Alignment of sequences and description of variation/polymorphisms among the aligned sequences, amounts of similarities and dissimilarities,
- Measurement of molecular variation among various regions of a gene, testing of neutrality models,
- Rates of nucleotide changes among coding and the non-coding regions within and among populations,
- Construction of phylogenetic trees using methods such as neighborhood joining and maximum parsimony; estimation of ages of variances using coalescent models,

***Population genetics***

- Patterns of distribution of genes among genotypes and populations. Hardy-Weinberg equilibrium, departures form the equilibrium
- Genotype and haplotype frequencies, levels of heterozygosities, polymorphism information contents of genes, estimation of haplotypes from genotypes; the E-M algorithm, and parsimony methods
- Estimation of linkage disequilibrium and recombination
- Hierarchical structure of populations, the F-statistics, estimation of inbreeding, selection and drift
- Genetic admixture/migration and mutation frequencies
- Spatial distribution of genotypes using spatial autocorrelation methods
- Kin-structured maintainance of variation and migration

*Quantitative genetics*

- Phenotype as the product of the interaction between genotype and environment
- Additive, dominance and epistatic variance on the phenotype
- Effects of homozygosity, heterozygosity and developmental homeostasis
- Estimation of heritability: broad sense and narrow sense

- Determination of number of genes governing a character
- Determination of quantitative trait loci (QTLs) using family information or population information, and using linkage and/or association studies
- Determination of quantitative trait nucleotide (QTN) using a combination linkage disequilibrium methods and cladistic approaches
- Determination of individual causal nucleotide in the diploid or haploid state on the phenotype using the method of measured genotype approaches, and combined effects or synergistic interaction of the causal mutations on the phenotype
- Determination of relative importance of each of the mutations on a given phenotype using multivariate methods, such as discriminant function, principal component and step-wise regression methods
- Determination of direct and indirect effect of polymorphisms on a complex phenotype using path analysis (partial regression ) methods
- Determination of the effects of specific environment on a given genotype – genotype x environment interactions using joint regression and additive and multiplicative parameter methods.

*Genetic epidemiology*

- Determination of sample size based on the disease and the marker frequency in the "case" and in the "control" populations
- Stratification of study population based on gender, ethnic, socio-economic variation
- Establishing a "causal relationship" between genotype and disease, using various association and linkage approaches – viz., case-control designs, family studies (if available), transmission disequilibrium tests etc.,
- Linkage analysis between markers and a candidate locus using two-point and multipoint approaches.

Computer programs used for genetic analysis are: Dna SP version 3.0, by Juilo Rozas, University of Barcelona, Spain. Http://www.bio.ub.es/-Julio; Arlequin 1.1 by S. Schnieder, J-M Kueffer, D. Roessli and L. Excoffier. University of Geneva, Switzerland, http://anthropologie.unige.ch/arlequin. PAUP*4, by D. L. Swofford, Sinauer Associates, Inc., 1999. SYSTAT software, SPSS Inc., Chicago, Il., 1998; . Linkage User's Guide, by J. Ott, Rockefeller University, Http://Linkage.rockefeller.edu/soft/linkage

Guidance in the selection of appropriate genetic statistical tests for analysis of data can be obtained from texts such as: Fundamentals of Genetic Epidemiology (Monographs in Epidemiology and Biostatistics, Vol 22) by M. J. Khoury, B. H. Cohen & T. H. Beaty, Oxford Univ Press, 1993; Methods in Genetic Epidemiology by Newton E. Morton, S. Karger Publishing, 1983; Methods in Observational Epidemiology, 2nd edition (Monographs in Epidemiology and Biostatistics, V. 26) by J. L. Kelsey (Editor), A. S. Whittemore & A. S. Evans, 1996; Clinical Trials : Design, Conduct, and Analysis (Monographs in Epidemiology and Biostatistics, Vol 8) by C. L. Meinert & S. Tonascia, 1986)

*I. Retrospective clinical trials.*

In general the goal of retrospective clinical trials is to test and refine hypotheses regarding genetic factors that are associated with drug responses. The best supported hypotheses can subsequently be tested in prospective clinical trials, and data from the prospective trials will likely comprise the main basis for an application to register the drug and predictive genetic test with the appropriate regulatory body. In some cases, however, it may become acceptable to use data from retrospective trials to support regulatory filings. Exemplary strategies and criteria for stratifying patients in a retrospective clinical trial are provided below.

Clinical trials to study the effect of one gene locus on drug response

A. Stratify patients by genotype at one candidate variance in the candidate gene locus.

1. Genetic stratification of patients can be accomplished in several ways, including the following (where 'A' is the more frequent form of the variance being assessed and 'a' is the less frequent form):

(a) AA vs. aa

(b) AA vs. Aa vs. aa

(c) AA vs. (Aa + aa)

(d) (AA + Aa) vs. aa.

2. The effect of genotype on drug response phenotype may be affected by a variety of nongenetic factors. Therefore it may be beneficial to measure the effect of genetic stratification in a subgroup of the overall clinical trial population. Subgroups can be defined in a number of ways including, for example, biological, clinical, pathological or environmental criteria. For example, the predictive value of genetic stratification can be assessed in a subgroup or subgroups defined by:

a. Biological criteria:

i. gender (males vs. females)

ii. age (for example above 60 years of age). Two, three or more age groups may be useful for defining subgroups for the genetic analysis.

iii. hormonal status and reproductive history, including pre- vs. post-menopausal status of women, or multiparous vs. nulliparous women

iv. ethnic, racial or geographic origin, or surrogate markers of ethnic, racial or geographic origin. (For a description of genetic markers that serve as surrogates of racial/thnic origin see, for example: Rannala, B. and J.L. Mountain, Detecting immigration by using multilocus genotypes. *Proc Natl Acad Sci U S A* , 94 (17):

9197-9201, 1997. Other surrogate markers could be used, including biochemical markers.)

b. Clinical criteria:

i. Disease status. There are clinical grading scales for many diseases. For example, the status of Alzheimer's Disease patients is often measured by cognitive assessment scales such as the mini-mental status exam (MMSE) or the Alzheimer's Disease Assessment Scale (ADAS), which includes a cognitive component (ADAS-COG). There are also clinical assessment scales for many other diseases, including cancer.

ii. Disease manifestations (clinical presentation).

iii. Radiological staging criteria.

c. Pathological criteria:

i. Histopathologic features of disease tissue, or pathological diagnosis. (For example there are many varieties of lung cancer: squamous cell carcinoma, adenocarcinoma, small cell carcinoma, bronchoalveolar carcinoma, etc., each of which may – which, in combination with genetic variation, may correlate with

ii. Pathological stage. A variety of diseases, particularly cancer, have pathological staging schemes

iii. Loss of heterozygosity (LOH)

iv. Pathology studies such as measuring levels of a marker protein

v. Laboratory studies such as hormone levels, protein levels, small molecule levels

3. Measure frequency of responders in each genetic subgroup. Subgroups may be defined in several ways.

i. more than two age groups

ii. reproductive status such as pre or post-menopausal

4. Stratify by haplotype at one candidate locus where the haplotype is made up of two variances, three variances or greater than three variances.

Data from already completed clinical trials can be retrospectively reanalyzed. Since the questions are new, the data can be treated as if it were a prospective trial, with identified variances or haplotypes as stratification criteria or endpoints in clinically stratified data (e.g. what is the frequency of a particular variance in a response group compared to nonresponsders). Care should be taken to in studying a population in which there may be a link between drug-related genes and disease-related genes.

Retrospective pharmacogenetic trials can be conducted at each of the phases

of clinical development, if sufficient data is available to correlate the physiologic effect of the candidate therapeutic intervention and the allelic variance or variances within the treatment population. In the case of a retrospective trial, the data collected from the trial can be re-analyzed by imposing the additional stratification on groups of patients by specific allelic variances that may exist in the treatment groups. Retrospective trials can be useful to ascertain whether a hypothesis that a specific variance has a significant effect on the efficacy or toxicity profile for a candidate therapeutic intervention.

A prospective clinical trial has the advantage that the trial can be designed to ensure the trial objectives can be met with statistical certainty. In these cases, power analysis, which includes the parameters of allelic variance frequency, number of treatment groups, and ability to detect positive outcomes can ensure that the trial objectives are met.

In designing a pharmacogenetic trial, retrospective analysis of Phase II or Phase III clinical data can indicate trial variables for which further analysis is beneficial. For example, surrogate endpoints, pharmacokinetic parameters, dosage, efficacy endpoints, ethnic and gender differences, and toxicological parameters may result in data that would require further analysis and re-examination through the design of an additional trial. In these cases, analysis involving statistics, genetics, clinical outcomes, and economic parameters may be considered prior to proceeding to the stage of designing any additional trials. Factors involved in the consideration of statistical significance may include Bonferroni analysis, permutation testing, with multiple testing correction resulting in a difference among the treatment groups that has occurred as a result of a chance of no greater than 20%, i.e. $p < 0.20$. Factors included in determining clinical outcomes to be relevant for additional testing may include, for example, consideration of the target indication, the trial endpoints, progression of the disease, disorder, or condition during the trial study period, biochemical or pathophysiologic relevance of the candidate therapeutic intervention, and other variables that were not included or anticipated in the initial study design or clinical protocol. Factors to be included in the economic significance in determining additional testing parameters include sample size, accrual rate, number of clinical sites or institutions required, additional or other available medical or therapeutic interventions approved for human use, and additional or other available medical or therapeutic interventions concurrently or anticipated to enter human clinical testing. Further, there may be patients within the treatment categories that present data that fall outside of the average or mean values, or there may be an indication of multiple allelic loci that are involved in the responses to the candidate therapeutic intervention. In these cases, one could propose a prospective clinical trial having an

objective to determine the significance of the variable or parameter and its effect on the outcome of the parent Phase II trial. In the case of a pharmacogenetic difference, i.e. a single or multiple allelic difference, a population could be selected based upon the distribution of genotypes. The candidate therapeutic intervention could then be tested in this group of volunteers to test for efficacy or toxicity. The repeat prospective study could be a Phase I limited study in which the subjects would be healthy human volunteers, or a Phase II limited efficacy study in which patients which satisfy the inclusion criteria could be enrolled. In either case, the second, confirmatory trial could then be used to systematically ensure an adequate number of patients with appropriate phenotype is enrolled in a Phase III trial.

A placebo controlled pharmacogenetics clinical trial design will be one in which target allelic variance or variances will be identified and a diagnostic test will be performed to stratify the patients based upon presence, absence, or combination thereof of these variances. In the Phase II or Phase III stage of clinical development, determination of a specific sample size of a prospective trial will be described to include factors such as expected differences between a placebo and treatment on the primary or secondary endpoints and a consideration of the allelic frequencies.

The design of a pharmacogenetics clinical trial will include a description of the allelic variance impact on the observed efficacy between the treatment groups. Using this type of design, the type of genetic and phenotypic relationship display of the efficacy response to a candidate therapeutic intervention will be analyzed. For example, a genotypically dominant allelic variance or variances will be those in which both heterozygotes and homozygotes will demonstrate a specific phenotypic efficacy response different from the homozygous recessive genotypic group. A pharmacogenetic approach is useful for clinicians and public health professionals to include or eliminate small groups of responders or non-responders from treatment in order to avoid unjustified side-effects. Further, adjustment of dosages when clear clinical difference between heterozygous and homozygous individuals may be beneficial for therapy with the candidate therapeutic intervention

In another example, a reccesive allelic variance or variances will be those in which only the homozygote recessive for that or those variances will demonstrate a specific phenotypic efficacy response different from the heterozygotes or homozygous dominants. An extension of these examples may include allelic variance or variances organized by haplotypes from additional gene or genes.

## V. Variance Identification and Use

A. Initial Identification of variances in genes

*Selection of population size and composition*

Prior to testing to identify the presence of sequence variances in a particular gene or genes, it is useful to understand how many individuals should be screened to provide confidence that most or nearly all pharmacogenetically relevant variances will be found. The answer depends on the frequencies of the phenotypes of interest and what assumptions we make about heterogeneity and magnitude of genetic effects. Prior to testing to identify the presence of sequence variances in a particular gene or genes, it is useful to understand how many individuals should be screened to provide confidence that most or nearly all pharmacogenetically relevant variances will be found. The answer depends on the frequencies of the phenotypes of interest and what assumptions we make about heterogeneity and magnitude of genetic effects. At the beginning we only know phenotype frequencies (e.g. responders vs. nonresponders, frequency of various side effects, etc.).

The most conservative assumption (resulting in the lowest estimate of allele frequency, and consequently the largest suggested screening population) is (i) that the phenotype (e.g. toxicity or efficacy) is multifactorial (i.e. can be caused by two or more variances or combinations of variances), (ii) that the variance of interest has a high degree of penetrance (i.e. is consistently associated with the phenotype), and (iii) that the mode of transmission is Mendelian dominant. Consider a pharmacogenetic study designed to identify predictors of efficacy for a compound that produces a 15% response rate in a nonstratified population. If half the response is sustantially attributable to a given variance, and the variance is consistently associated with a positive response (in 80% of cases) and the variance need only be present in one copy to produce a positive result then ~10% of the subjects are likely heterozygotes for the variance that produces the response. The Hardy-Weinberg equation can be used to infer an allele frequency in the range of 5% from these assumptions (given allele frequencies of 5%/95% then: 2 x .05 x .95 = .095, or 9.5% heterozygotes are expected, and 0.05 x 0.05 = 0.0025, or 0.25% homozygotes are expected. They sum to 9.5% + 0.25% = 9.75% likely responders, 80% of whom, or 7.6%, are likely real responders due to presence of the positive response allele. Thus about half of the 15% responders are accounted for.). From the Table it can be seen that, in order to have a 99% chance of detecting an allele present at a frequency of 5% nearly 50 subjects should be screened for variances, assuming that the variances occur in the screening population at the same frequency as they occur in the patient population. Similar analyses can be performed for other assumptions regarding likely magnitude of effect, penetrance and mode of genetic transmission.

At the beginning we only know phenotype frequencies (e.g. responders vs. nonresponders, frequency of various side effects, etc.). As an example, the occurrence of serious 5-FU/FA toxicity - e.g. toxicity requiring hospitalization is often >10%. The occurrence of life threatening toxicity is in the 1-3% range (Buroker et al. 1994). The occurrence of complete remissions is on the order of 2-8%. The lowest frequency phenotypes are thus on the order of ~2%. If we assume that (i) homogeneous genetic effects are responsible for half the phenotypes of interest and (ii) for the most part the extreme phenotypes represent recessive genotypes, then we need to detect alleles that will be present at ~10% frequency (.1 x .1 = .01, or 1% frequency of homozygotes) if the population is at Hardy-Weinberg equilibrium. To have a ~99% chance of identifying such alleles would require searching a population of 22 individuals (see Table below). If the major phenotypes are associated with heterozygous genotypes then we need to detect alleles present at ~.5% frequency (2 x .005 x .995 = .00995, or ~1% frequency of heterozygotes). A 99% chance of detecting such alleles would require ~40 individuals (Table below). Given the heterogeneity of the North American population we cannot assume that all genotypes are present in Hardy-Weinberg proportions, therefore a substantial oversampling may be done to increase the chances of detecting relevant variances: For our initial screening, usually, 62 individuals of known race/ethnicity are screened for variance. Variance detection studies can be extended to outliers for the phenotypes of interest to cover the possibility that important variances were missed in the normal population screening.

| | Number of subjects genotyped | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Allele frequencies | n = 5 | n = 10 | n = 15 | n = 20 | n = 25 | n = 30 | n = 35 | n = 50 |
| p=.99, | 9.56 | 18.21 | 26.03 | 33.10 | 39.50 | 45.28 | 50.52 | 63.40 |
| p=.97, | 26.26 | 45.62 | 59.90 | 70.43 | 78.19 | 83.92 | 88.14 | 95.24 |
| p=.95, | 40.13 | 64.15 | 78.53 | 87.15 | 92.30 | 95.39 | 97.24 | 99.65 |
| p=.93, | 51.60 | 76.58 | 88.66 | 94.51 | 97.34 | 98.71 | 99.38 | 99.93 |
| p =.9, q = | 65.13 | 87.84 | 95.76 | 98.52 | 99.48 | 99.82 | 99.94 | >99.9 |
| p =.8, q = | 89.26 | 98.84 | 99.88 | 99.99 | >99.9 | >99.9 | >99.9 | >99.9 |
| p =.7, q = | 97.17 | 99.92 | 99.99 | >99.9 | >99.9 | >99.9 | >99.9 | >99.9 |

*Likelihood of Detecting Polymorphism in a Population as a Function of Allele Frequency & Number of Individuals Genotyped*

The table above shows the probability (expressed as percent) of detecting both alleles (i.e. detecting heterozygotes) at a biallelic locus as a function of (i) the allele frequencies and (ii) the number of individuals genotyped. The chances of detecting heterozygotes increases as the frequencies of the two alleles approach 0.5 (down a column), and as the number of individuals genotyped increases (to the right along a row). The numbers in the table are given by the formula: $1 - (p)^{2n} - (q)^{2n}$. Allele frequencies are designated p and q and the number of individuals tested is designated n. (Since humans are diploid, the number of alleles tested is twice the number of individuals, or 2n.)

While it is preferable that numbers of individuals, or independent sequence samples, are screened to identify variances in a gene, it is also very beneficial to identify variances using smaller numbers of individuals or sequence samples. For example, even a comparison between the sequences of two samples or individuals can reveal sequence variances between them. Preferably, 5, 10, or more samples or individuals are screened.

*Source of nucleic acid samples*

Nucleic acid samples, for example for use in variance identification, can be obtained from a variety of sources as known to those skilled in the art, or can be obtained from genomic or cDNA sources by known methods. For example, the Coriell Cell Repository (Camden, N.J.) maintains over 6,000 human cell cultures, mostly fibroblast and lymphoblast cell lines comprising the NIGMS Human Genetic Mutant Cell Repository. A catalog (http://locus.umdnj.edu/nigms) provides racial or ethnic identifiers for many of the cell lines. It is preferable to perform polymorphism discovery on a population that mimics the population to be evaluated in a clinical trial, both in terms of racial/ethnic/geographic background and in terms of disease status. Otherwise, it is generally preferable to include a broad population sample including, for example, (for trials in the United States): Caucasians of Northern, Central and Southern European origin, Africans or African-Americans, Hispanics or Mexicans, Chinese, Japanese, American Indian, East Indian, Arabs and Koreans.

*Source of human DNA, RNA and cDNA samples*

PCR based screening for DNA polymorphism can be carried out using either genomic DNA or cDNA produced from mRNA. For many genes, only cDNA sequences have been published, therefore the analysis of those genes is, at least initially, at the cDNA level since the determination of intron-exon boundaries and the isolation of flanking sequences is a laborious process. However, screening

genomic DNA has the advantage that variances can be identified in promoter, intron and flanking regions. Such variances may be biologically relevant. Therefore preferably, when variance analysis of patients with outlier responses is performed, analysis of selected loci at the genomic level is also performed. Such analysis would be contingent on the availability of a genomic sequence or intron-exon boundary sequences, and would also depend on the anticipated biological importance of the gene in connection with the particular response.

When cDNA is to be analyzed it is very beneficial to establish a tissue source in which the genes of interest are expressed at sufficient levels that cDNA can be readily produced by RT-PCR. Preliminary PCR optimization efforts for 19 of the 29 genes in Table 2 reveal that all 19 can be amplified from lymphoblastoid cell mRNA. The 7 untested genes belong on the same pathways and are expected to also be PCR amplifiable.

*PCR Optimization*

Primers for amplifying a particular sequence can be designed by methods known to those skilled in the art, including by the use of computer programs such as the PRIMER software available from Whitehead Institute/MIT Genome Center. In some cases it is preferable to optimize the amplification process according to parameters and methods known to those skilled in the art; optimization of PCR reactions based on a limited array of temperature, buffer and primer concentration conditions is utilized. New primers are obtained if optimization fails with a particular primer set.

*Variance detection using T4 endonuclease VII mismatch cleavage method*

Any of a variety of different methods for detecting variances in a particular gene can be utilized, such as those described in the patents and applications cited in section A above. An exemplary method is a T4 EndoVII method. The enzyme T4 endonuclease VII (T4E7) is derived from the bacteriophage T4. T4E7 specifically cleaves heteroduplex DNA containing single base mismatches, deletions or insertions. The site of cleavage is 1 to 6 nucleotides 3' of the mismatch. This activity has been exploited to develop a general method for detecting DNA sequence variances (Youil et al. 1995; Mashal and Sklar, 1995). A quality controlled T4E7 variance detection procedure based on the T4E7 patent of R.G.H. Cotton and co-workers. (Del Tito et al., in press) is preferably utilized. T4E7 has the advantages of being rapid, inexpensive, sensitive and selective. Further, since the enzyme

pinpoints the site of sequence variation, sequencing effort can be confined to a 25 - 30 nucleotide segment.

The major steps in identifying sequence variations in candidate genes using T4E7 are: (1) PCR amplify 400-600 bp segments from a panel of DNA samples; (2) mix a fluorescently-labeled probe DNA with the sample DNA; (3) heat and cool the samples to allow the formation of heteroduplexes; (4) add T4E7 enzyme to the samples and incubate for 30 minutes at 37°C, during which cleavage occurs at sequence variance mismatches; (5) run the samples on an ABI 377 sequencing apparatus to identify cleavage bands, which indicate the presence and location of variances in the sequence; (6) a subset of PCR fragments showing cleavage are sequenced to identify the exact location and identity of each variance.

The T4E7 Variance Imaging procedure has been used to screen particular genes. The efficiency of the T4E7 enzyme to recognize and cleave at all mismatches has been tested and reported in the literature. One group reported detection of 81 of 81 known mutations (Youil et al. 1995) while another group reported detection of 16 of 17 known mutations (Mashal and Sklar, 1995). Thus, the T4E7 method provides highly efficient variance detection.

*DNA sequencing*

A subset of the samples containing each unique T4E7 cleavage site is selected for sequencing. DNA sequencing can, for example, be performed on ABI 377 automated DNA sequencers using BigDye chemistry and cycle sequencing. Analysis of the sequencing runs will be limited to the 30-40 bases pinpointed by the T4E7 procedure as containing the variance. This provides the rapid identification of the altered base or bases.

In some cases, the presence of variances can be inferred from published articles which describe Restriction Fragment Length Polymorphisms (RFLP). The sequence variances or polymorphisms creating those RFLPs can be readily determined using convention techniques, for example in the following manner. If the RFLP was initially discovered by the hybridization of a cDNA, then the molecular sequence of the RFLP can be determined by restricting the cDNA probe into fragments and separately hybridizing to a Southern blot consisting of the restriction digestion with the enzyme which reveals the polymorphic site, identifying the sub-fragment which hybridizes to the polymorphic restriction fragment, obtaining a genomic clone of the gene (e.g., from commercial services such as Genome Systems (Saint Louis, Missouri) or Research Genetics (Alabama) which will provide appropriate genomic clones on receipt of appropriate primer pairs). Using the genomic clone, restrict the genomic clone with the restriction enzyme

which revealed the polymorphism and isolate the fragment which contains the polymorphism, e.g., identifying by hybridization to the cDNA which detected the polymorphism. The fragment is then sequenced across the polymorphic site. A copy of the other allele can be obtained by PCT from addition samples.

*Variance detection using sequence scanning*

In addition to the physical methods, e.g., those described above and others known to those skilled in the art (see, e.g., Housman, U.S. Patent 5,702,890; Housman et al., U.S. Patent Application 09/045,053), variances can be detected using computational methods, involving computer comparison of sequences from two or more different biological sources, which can be obtained in various ways, for example from public sequence databases. The term "variance scanning" refers to a process of identifying sequence variances using computer-based comparison and analysis of multiple representations of at least a portion of one or more genes. Computational variance detection involves a process to distinguish true variances from sequencing errors or other artifacts, and thus does not require perfectly accurate sequences. Such scanning can be performed in a variety of ways, preferably, for example, as described in Stanton et al., filed October 14, 1999, serial number 09/419,705, attorney docket number 246/128.

While the utilization of complete cDNA sequences is highly preferred, it is also possible to utilize genomic sequences. Such analysis may be desired where the detection of variances in or near splice sites is sought. Such sequences may represent full or partial genomic DNA sequences for a gene or genes. Also, as previously indicated, partial cDNA sequences can also be utilized although this is less preferred. As described below, the variance scanning analysis can simply utilize sequence overlap regions, even from partial sequences. Also, while the present description is provided by reference to DNA, e.g., cDNA, some sequences may be provided as RNA sequences, e.g., mRNA sequences. Such RNA sequences may be converted to the corresponding DNA sequences, or the analysis may use the RNA sequences directly.

B. Determination of Presence or Absence of Known Variances

The identification of the presence of previously identified variances in cells of an individual, usually a particular patient, can be performed by a number of different techniques as indicated in the Summary above. Such methods include methods utilizing a probe which specifically recognizes the presence of a particular nucleic acid or amino acid sequence in a sample. Common types of probes include nucleic acid hybridization probes and antibodies, for example, monoclonal

antibodies, which can differentially bind to nucleic acid sequences differing in one or more variance sites or to polypeptides which differ in one or more amino acid residues as a result of the nucleic acid sequence variance or variances. Generation and use of such probes is well-known in the art and so is not described in detail herein.

Preferably, however, the presence or absence of a variance is determined using nucleotide sequencing of a short sequence spanning a previously identified variance site. This will utilize validated genotyping assays for the polymorphisms previously identified. Since both normal and tumor cell genotypes can be measured, and since tumor material will frequently only be available as paraffin embedded sections (from which RNA cannot be isolated), it will be necessary to utilize genotyping assays that will work on genomic DNA. Thus PCR reactions will be designed, optimized, and validated to accommodate the intron-exon structure of each of the genes. If the gene structure has been published (as it has for some of the listed genes), PCR primers can be designed directly. However, if the gene structure is unknown, the PCR primers may need to be moved around in order to both span the variance and avoid exon-intron boundaries. In some cases one-sided PCR methods such as bubble PCR (Ausubel et al. 1997) may be useful to obtain flanking intronic DNA for sequence analysis.

Using such amplification procedures, the standard method used to genotype normal and tumor tissues will be DNA sequencing. PCR fragments encompassing the variances will be cycle sequenced on ABI 377 automated sequencers using Big Dye chemistry

C. Correlation of the Presence or Absence of Specific Variances with Differential Treatment Response

Prior to establishment of a diagnostic test for use in the selection of a treatment method or elimination of a treatment method, the presence or absence of one or more specific variances in a gene or in multiple genes is correlated with a differential treatment response. (As discussed above, usually the existence of a variable response and the correlation of such a response to a particular gene is performed first.) Such a differential response can be determined using prospective and/or retrospective data. Thus, in some cases, published reports will indicate that the course of treatment will vary depending on the presence or absence of particular variances. That information can be utilized to create a diagnostic test and/or incorporated in a treatment method as an efficacy or safety determination step.

Usually, however, the effect of one or more variances is separately determined. The determination can be performed by analyzing the presence or

absence of particular variances in patients who have previously been treated with a particular treatment method, and correlating the variance presence or absence with the observed course, outcome, and/or development of adverse events in those patients. This approach is useful in cases in which observation of treatment effects was clearly recorded and cell samples are available or can be obtained. Alternatively, the analysis can be performed prospectively, where the presence or absence of the variance or variances in an individual is determined and the course, outcome, and/or development of adverse events in those patients is subsequently or concurrently observed and then correlated with the variance determination.

*Analysis of Haplotypes Increases Power of Genetic Analysis*

In some cases, variation in activity due to a single gene or a single genetic variance in a single gene may not be sufficient to account for a clinically significant fraction of the observed variation in patient response to a treatment, e.g., a drug, there may be other factors that account for some of the variation in patient response. Drug response phenotypes may vary continuously, and such (quantitative) traits may be influenced by a number of genes (Falconer and Mackay, Quantitative Genetics, 1997). Although it is impossible to determine *a priori* the number of genes influencing a quantitative trait, potentially only one or a few loci have large effects, where a large effect is 5-20% of total variation in the phenotype (Mackay, 1995).

Having identified genetic variation in enzymes that may affect action of a specific drug, it is useful to efficiently address its relation to phenotypic variation. The sequential testing for correlation between phenotypes of interest and single nucleotide polymorphisms may be adequate to detect associations if there are major effects associated with single nucleotide changes; certainly it is useful to this type of analysis. However there is no way to know in advance whether there are major phenotypic effects associated with single nucleotide changes and, even if there are, there is no way to be sure that the salient variance has been identified by screening cDNAs. A more powerful way to address the question of genotype-phenotype correlation is to assort genotypes into haplotypes. (A haplotype is the cis arrangement of polymorphic nucleotides on a particular chromosome.) Haplotype analysis has several advantages compared to the serial analysis of individual polymorphisms at a locus with multiple polymorphic sites.

(1) Of all the possible haplotypes at a locus ($2^n$ haplotypes are theoretically possible at a locus with n binary polymorphic sites) only a small fraction will

generally occur at a significant frequency in human populations. Thus, association studies of haplotypes and phenotypes will involve testing fewer hypotheses. As a result there is a smaller probability of Type I errors, that is, false inferences that a particular variant is associated with a given phenotype.

(2) The biological effect of each variance at a locus may be different both in magnitude and direction. For example, a polymorphism in the 5' UTR may affect translational efficiency, a coding sequence polymorphism may affect protein activity, a polymorphism in the 3' UTR may affect mRNA folding and half life, and so on. Further, there may be interactions between variances: two neighboring polymorphic amino acids in the same domain - say cys/arg at residue 29 and met/val at residue 166 - may, when combined in one sequence, for example, 29cys-166val, have a deleterious effect, whereas 29cys-166met, 29arg-166met and 29arg-166val proteins may be nearly equal in activity. Haplotype analysis is the best method for assessing the interaction of variances at a locus.

(3) Templeton and colleagues have developed powerful methods for assorting haplotypes and analyzing haplotype/phenotype associations (Templeton et al., 1987). Alleles which share common ancestry are arranged into a tree structure (cladogram) according to their (inferred) time of origin in a population (that is, according to the principle of parsimony). Haplotypes that are evolutionarily ancient will be at the center of the branching structure and new ones (reflecting recent mutations) will be represented at the periphery, with the links representing intermediate steps in evolution. The cladogram defines which haplotype-phenotype association tests should be performed to most efficiently exploit the available degrees of freedom, focusing attention on those comparisons most likely to define functionally different haplotypes (Haviland et al., 1995). This type of analysis has been used to define interactions between heart disease and the apolipoprotein gene cluster (Haviland et al 1995) and Alzheimer's Disease and the Apo-E locus (Templeton 1995) among other studies, using populations as small as 50 to 100 individuals. The methods of Templeton have also been applied to meaure the genetic determinants of variation in the angiotensin-I converting enzyme gene. (Keavney, B., McKenzie, C. A., Connoll, J.M.C., et al. Measured haplotype analysis of the angiotensin-I converting enzyme gene. *Human Molecular Genetics* 7: 1745-1751.)

*Methods for determining haplotypes*

The goal of haplotyping is to identify the common haplotypes at selected loci that have multiple sites of variance. Haplotypes are usually determined at the cDNA level. Several general approaches to identification of haplotyes can be employed. Haplotypes may also be estimated using computational methods or determined definitively using experimental approaches. Computational approachs generally include an expectation maximization (E-M) algorithm (see, for example: Excoffier and Slatkin, Mol. Biol. Evol. 1995) or a combination of Parsimony (see below) and E-M methods.

Haplotypes can be determined experimentally without requirement of a haplotyping method by genotyping samples from a set of pedigrees and observing the segregation of haplotypes. For example families collected by the Centre d'Etude du Polymorphisme Humaine (CEPH) can be used. Cell lines from these families are available from the Coriell Repository. This approach will be useful for cataloging common haplotypes and for validating methods on samples with known haplotypes. The set of haplotypes determined by pedigree analysis can be useful in computational methods, including those utilizing the E-M algorithm.

Haplotypes can also be determined directly from cDNA using the T4E7 procedure. T4E7 cleaves mismatched heteroduplex DNA at the site of the mismatch. If a heteroduplex contains only one mismatch, cleavage will result in the generation of two fragments. However, if a single heteroduplex (allele) contains two mismatches, cleavage will occur at two different sites resulting in the generation of three fragments. The appearance of a fragment whose size corresponds to the distance between the two cleavage sites is diagnostic of the two mismatches being present on the same strand (allele). Thus, T4E7 can be used to determine haplotypes in diploid cells.

An alternative method, allele specific PCR, may be used for haplotyping. The utility of allele specific PCR for haplotyping has already been established (Michalatos-Beloin et al., 1996; Chang et al. 1997). Opposing PCR primers are designed to cover two sites of variance (either adjacent sites or sites spanning one or more internal variances). Two versions of each primer are synthesized, identical to each other except for the 3' terminal nucleotide. The 3' terminal nucleotide is designed so that it will hybridize to one but not the other variant base. PCR amplification is then attempted with all four possible primer combinations in separate wells. Because Taq polymerase is very inefficient at extending 3' mismatches, the only samples which will be amplified will be the ones in which the two primers are perfectly matched for sequences on the same strand (allele). The presence or absence of PCR product allows haplotyping of diploid cell lines. At

most two of four possible reactions should yield products. This procedure has been successfully applied, for example, to haplotype the DPD amino acid polymorphisms.

Parsimony methods are also useful for classifying DNA sequences, haplotypes or phenotypic characters. Parsimony principle maintains that the best explanation for the observed differences among sequences, phenotypes (individuals, species) etc., is provided by the smallest number of evolutionary changes. Alternatively, simpler hypotheses are preferable to explain a set of data or patterns, than more complicated ones, and *ad hoc* hypotheses should be avoided whenever possible (Molecular Systematics, Hillis et al., 1996). Parsimony methods thus operate by minimizing the number of evolutionary steps or mutations (changes from one sequence/character) required to account for a given set of data.

For example, supposing we want to obtain relationships among a set of sequences and construct a structure (tree/topology), we first count the minimum number of mutations that are required for explaining the observed evolutionary changes among a set of sequences. A structure (topology) is constructed based on this number. When once this number is obtained, another structure is tried. This process is continued for all reasonable number of structures. Finally, the structure that required the smallest number of mutational steps is chosen as the likely structure/evolutionary tree for the sequences studied.

D. Selection of Treatment Method Using Variance Information

1. General

Once the presence or absence of a variance or variances in a gene or genes is shown to correlate with the efficacy or safety of a treatment method, that information can be used to select an appropriate treatment method for a particular patient. In the case of a treatment which is more likely to be effective when administered to a patient who has at least one copy of a gene with a particular variance or variances (in some cases the correlation with effective treatment is for patients who are homozygous for a variance or set of variances in a gene) than in patients with a different variance or set of variances, a method of treatment is selected (and/or a method of administration) which correlates positively with the particular variance presence or absence which provides the indication of effectiveness. As indicated in the Summary, such selection can involve a variety of different choices, and the correlation can involve a variety of different types of treatments, or choices of methods of treatment. In some cases, the selection may include choices between treatments or methods of administration where more than one method is likely to be effective, or where there is a range of expected effectiveness or different expected levels of contra-indication or deleterious effects.

In such cases the selection is preferably performed to select a treatment which will be as effective or more effective than other methods, while having a comparatively low level of deleterious effects. Similarly, where the selection is between method with differing levels of deleterious effects, preferably a method is selected which has low such effects but which is expected to be effective in the patient.

Alternatively, in cases where the presence or absence of the particular variance or variances is indicative that a treatment or method of administration is more likely to be ineffective or contra-indicated in a patient with that variance or variances, then such treatment or method of administration is generally eliminated for use in that patient.

2. Diagnostic Methods

Once a correlation between the presence and absence of at least one variance in a gene or genes and an indication of the effectiveness of a treatment, the determination of the presence or absence of that at least one variance provides diagnostic methods, which can be used as indicated in the Summary above to select methods of treatment, methods of administration of a treatment, methods of selecting a patient or patients for a treatment and others aspects in which the determination of the presence or absence of those variances provides useful information for selecting or designing or preparing methods or materials for medical use in the aspects of this invention. As previously stated, such variance determination or diagnostic methods can be performed in various ways as understood by those skilled in the art.

In certain variance determination methods, it is necessary or advantageous to amplify one or more nucleotide sequences in one or more of the genes identified herein. Such amplification can be performed by conventional methods, e.g., using polymerase chain reaction (PCR) amplification. Such amplification methods are well-known to those skilled in the art and will not be specifically described herein. For most applications relevant to the present invention, a sequence to be amplified includes at least one variance site, which is preferably a site or sites which provide variance information indicative of the effectiveness of a method of treatment or method of administration of a treatment, or effectiveness of a second method of treatment which reduces a deleterious effect of a first treatment method, or which enhances the effectiveness of a first method of treatment. Thus, for PCR, such amplification generally utilizes primer oligonucleotides which bind to or extent through at least one such variance site under amplification conditions.

For convenient use of the amplified sequence, e.g., for sequencing, it is beneficial that the amplified sequence be of limited length, but still long enough to

allow convenient and specific amplification. Thus, preferably the amplified sequence has a length as described in the Summary.

Also, in certain variance determination, it is useful to sequence one or more portions of a gene or genes, in particular, portions of the genes identified in this disclosure. As understood by persons familiar with nucleic acid sequencing, there are a variety of effective methods. In particular, sequencing can utilize dye termination methods and mass spectrometric methods. The sequencing generally involves a nucleic acid sequence which includes a variance site as indicated above in connection with amplification. Such sequencing can directly provide determination of the presence or absence of a particular variance or set of variances, e.g., a haplotype, by inspection of the sequence (visually or by computer). Such sequencing is generally conducted on PCR amplified sequences in order to provide sufficient signal for practical or reliable sequence determination.

Likewise, in certain variance determinations, it is useful to utilize a probe or probes. As previously described, such probes can be of a variety of different types.

**VII.** Loss of Heterozygosity and Conditionally Essential Genes

Different environmental, pharmacological, and physical changes in the environment that result in homeostatic or compensatory responses in which genes that are not normally essential for cell survival or proliferation become essential are known in the art.

When LOH results in a difference in normal cell genotype vs. cancer cell genotype that affects a locus encoding a product affecting the cells' ability to survive in the presence of an environmental change, or a pharmaceutical or biological agent, or a physical factor, there is an opportunity to exploit a therapeutic window between cancer cells and normal cells. Below we describe specific examples of genes that (1) affect cell responses to altered environments, (2) are located on chromosomes that undergo LOH in cancer and (3) exist in two or more variant forms. These examples have been selected to illustrate how the therapeutic strategy described in this application would work with a variety of different alterations in chemical or physical environment. Example 20 describes a gene (Dihydropyrimidine Dehydrogenase) that mediates response to an altered chemical environment (presence of the toxic chemical 5-floxuridine) by specifically transforming the chemical to an inactive metabolite. Example 27 describes a gene (Methylguanine methyltransferase) that mediates response to an altered chemical environment (presence of toxic chemicals such as nitrosourea or other alkylating agents) by removing methyl or alkyl adducts to DNA, the principal toxic lesion of these agents.

Example 21 describes a set of genes (Fanconi Anemia genes A,B,C,D,E,F,G and H) which mediate response to an altered chemical environment (presence of chemicals which cause DNA crosslinking, such as diepoxybutane, mitomycin C and cisplatinum) by repairing the crosslinks. Example 25 describes a set of genes (the DNA Dependent Protein Kinase Complex, including the DNA Dependent Protein Kinase catalytic subunit (DNA-PKcs), the DNA binding component (called Ku), made up of Ku-70 and Ku-86 kDa subunits, and the Ku-86 related protein Karp-1) that mediates repair of double stranded DNA breaks, such as occurs after x-irradiation. Example 22 describes a gene (asparagine synthase) that mediates response to an altered nutritional environment (absence of extracellular asparagine) which can be produced by an enzyme such as asparaginase, which hydrolyzes serum asparagine. Example 26 describes the Ataxia Telangiectasia gene, which is involved in response to ionizing radiation and radiomimetic chemicals. Other detailed examples include methionine synthase (Ex. 23) and methylthioadenosine phosphorylase (Ex. 24). Other examples include Poly (ADP) Ribose Polymerase (PARP), Glutathione-S- Transferase pi (GST-pi), NF-kappa B, Abl Kinase, 3-alkaylguanine alkyltransferase, N-methylpurine DNA glycosylase (hydrolyzes the deoxyribose N-glycosidic bond to excise 3-methyladenine and 7-methylguanine from alkylating agent-damaged DNA polymers), OGG-1, MDR-1.

In addition to the direct use of conditionally essential (or essential) genes in allele-specific inhibitor applications, the information provided by the LOH status of a gene. For example, in some cases, the effect of LOH can be a gene dosage effeect. This can additionally be combined with a reduced activity associated with particular forms of the gene. Either or both types of information can be used to identify patients who would be expected to respond differently to a treatment targeting that gene than would patients with two copies of the gene, or with at least one copy of a different form of the gene than remained after LOH. To illustrate, a patient may be heterozygous for a high activity allele and a low activity allele. LOH in cancer cells could remove either the high activity allele or the low activity allele, leaving only the other allele in cancer cells in the patient, while the normal cells would have intermediate activity due the presenece of both alleles. As a result, a therapy targeting or otherwise involving that gene in the response to treatment would be expected to result in variation in response between the normal cells and the cancer cells in the patient. If the low activity allele correlated with high response to the therapy, then it would be expected that the anti-cancer treatment would be more effective in a patient with such LOH than in a patient in whom cancer cells had not undergone LOH with respect to that gene.

Indeed, LOH assays for particular genes can also be used as surrogate assays for other LOH of other genes located near the marker gene. Thus, the marker gene can, for example, be used in connection with LOH-related effects or evaluations of other nearby genes. Such genes can include genes in the same pathway, as those genes are often located in close proximity on the same chromosome.

In has been shown that LOH at tumor suppressor genes correlates with anticancer chemotherapy response. Thus, LOH information on tumor suppressor genes can also be used in connection with LOH and/or pharmacogenetic information about other genes. As a result, it is beneficial to determine both the LOH status of the tumor suppressor gene or genes and one or more additional genes.

Together, or separately, the LOH information and the variance-based pharmacogenetic information can be used to identifiy patient subset that will respond differently to a particular therapy related to particular genes and/or to select appropriate therapies for patients based on the forms of the gene or genes in disease cells and normal cells.

## VII. Pharmaceutical Compositions, Including Pharmaceutical Compositions Adapted to be Preferentially Effective in Patients Having Particular Genetic Characteristics

A. General

The methods of the present invention, in many cases will utilize conventional pharmaceutical compositions, but will allow more advantageous and beneficial use of those compositions due to the ability to identify patients who are likely to benefit from a particular treatment or to identify patients for whom a particular treatment is less likely to be effective or for whom a particular treatment is likely to produce undesirable or intolerable effects. However, in some cases, it is advantageous to utilize compositions which are adapted to be preferentially effective in patients who possess particular genetic characteristics, i.e., in whom a particular variance or variances in one or more genes is present or absent (depending on whether the presence or the absence of the variance or variances in a patient is correlated with an increased expectation of beneficial response). Thus, for example, the presence of a particular variance or variances may indicate that a patient can beneficially receive a significantly higher dosage of a drug than a patient having a different

B. Regulatory Indications and Restrictions

The sale and use of drugs and the use of other treatment methods usually are subject to certain restrictions by a government regulatory agency charged with

ensuring the safety and efficacy of drugs and treatment methods for medical use, and approval is based on particular indications. In the present invention it is found that variability in patient response or patient tolerance of a drug or other treatment often correlates with the presence or absence of particular variances in particular genes. Thus, it is expected that such a regulatory agency may indicate that the approved indications for use of a drug with a variance-related variable response or toleration include use only in patients in whom the drug will be effective, and/or for whom the administration of the drug will not have intolerable deleterious effects, such as excessive toxicity or unacceptable side-effects. Conversely, the drug may be given for an indication that it may be used in the treatment of a particular disease or condition where the patient has at least one copy of a particular variance, variances, or variant form of a gene. Even if the approved indications are not narrowed to such groups, the regulatory agency may suggest use limited to particular groups or excluding particular groups or may state advantages of use or exclusion of such groups or may state a warning on the use of the drug in certain groups. Consistent with such suggestions and indications, such an agency may suggest or recommend the use of a diagnostic test to identify the presence or absence of the relevant variances in the prospective patient. Such diagnostic methods are described in this description. Generally, such regulatory suggestion or indication is provided in a product insert or label, and is generally reproduced in references such as the Physician's Desk Reference (PDR). Thus, this invention also includes drugs or pharmaceutical compositions which carry such a suggestion or statement of indication or warning or suggestion for a diagnostic test, and which may also be packaged with an insert or label stating the suggestion or indication or warning or suggestion for a diagnostic test.

In accord with the possible variable treatment responses, an indication or suggestion can specify that a patient be heterozygous, or alternatively, homozygous for a particular variance or variances or variant form of a gene. Alternatively, an indication or suggestion may specify that a patient have no more than one copy, or zero copies, of a particular variance, variances, or variant form of a gene.

A regulatory indication or suggestion may concern the variances or variant forms of a gene in normal cells of a patient and/or in cells involved in the disease or condition. For example, in the case of a cancer treatment, the response of the cancer cells can depend on the form of a gene remaining in cancer cells following loss of heterozygosity affecting that gene. Thus, even though normal cells of the patient may contain a form of the gene which correlates with effective treatment response, the absence of that form in cancer cells will mean that the treatment would be less likely to be effective in that patient than in another patient who retained in cancer

cells the form of the gene which correlated with effective treatment response. Those skilled in the art will understand whether the variances or gene forms in normal or disease cells are most indicative of the expected treatment response, and will generally utilize a diagnostic test with respect to the appropriate cells. Such a cell type indication or suggestion may also be contained in a regulatory statement, e.g., on a label or in a product insert.

C. Preparation and Administration of Drugs and Pharmaceutical Compositions Including Pharmaceutical Compositions Adapted to be Preferentially Effective in Patients Having Particular Genetic Characteristics

A particular compound useful in this invention can be administered to a patient either by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s). In treating a patient exhibiting a disorder of interest, a therapeutically effective amount of a agent or agents such as these is administered. A therapeutically effective dose refers to that amount of the compound that results in amelioration of one or more symptoms or a prolongation of survival in a patient.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the $LD_{50}$ (the dose lethal to 50% of the population) and the $ED_{50}$ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio $LD_{50}/ED_{50}$. Compounds which exhibit large therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the $ED_{50}$ with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized.

For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. For example, a dose can be formulated in animal models to achieve a circulating plasma concentration range that includes the $IC_{50}$ as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by HPLC.

The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. (See *e.g.* Fingl et. al., in The Pharmacological Basis of Therapeutics, 1975, Ch. 1 p.1). It should be noted that the attending physician would know how to and when to terminate, interrupt, or adjust administration due to toxicity, or to organ dysfunctions. Conversely, the

attending physician would also know to adjust treatment to higher levels if the clinical response were not adequate (precluding toxicity). The magnitude of an administrated dose in the management of disorder of interest will vary with the severity of the condition to be treated and the route of administration. The severity of the condition may, for example, be evaluated, in part, by standard prognostic evaluation methods. Further, the dose and perhaps dose frequency, will also vary according to the age, body weight, and response of the individual patient. A program comparable to that discussed above may be used in veterinary medicine.

Depending on the specific conditions being treated, such agents may be formulated and administered systemically or locally. Techniques for formulation and administration may be found in Remington's Pharmaceutical Sciences, 18th ed., Mack Publishing Co., Easton, PA (1990). Suitable routes may include oral, rectal, transdermal, vaginal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections, just to name a few.

For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For such transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

Use of pharmaceutically acceptable carriers to formulate the compounds herein disclosed for the practice of the invention into dosages suitable for systemic administration is within the scope of the invention. With proper choice of carrier and suitable manufacturing practice, the compositions of the present invention, in particular, those formulated as solutions, may be administered parenterally, such as by intravenous injection. The compounds can be formulated readily using pharmaceutically acceptable carriers well known in the art into dosages suitable for oral administration. Such carriers enable the compounds of the invention to be formulated as tablets, pills, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated.

Agents intended to be administered intracellularly may be administered using techniques well known to those of ordinary skill in the art. For example, such agents may be encapsulated into liposomes, then administered as described above. Liposomes are spherical lipid bilayers with aqueous interiors. All molecules present in an aqueous solution at the time of liposome formation are incorporated into the aqueous interior. The liposomal contents are both protected from the external microenvironment and, because liposomes fuse with cell membranes, are efficiently

delivered into the cell cytoplasm. Additionally, due to their hydrophobicity, small organic molecules may be directly administered intracellularly.

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. Determination of the effective amounts is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. The preparations formulated for oral administration may be in the form of tablets, dragees, capsules, or solutions. The pharmaceutical compositions of the present invention may be manufactured in a manner that is itself known, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levitating, emulsifying, encapsulating, entrapping or lyophilizing processes.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

Pharmaceutical preparations for oral use can be obtained by combining the active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and

suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added.

The invention described herein features methods for determining the appropriate identification of a patient diagnosed with a neurolgical disease or neurolgical dysfunction based on an analysis of the patient's allele status for a gene listed in Tables 1-6, 12-17, and 18-23. Specifically, the presence of at least one allele indicates that a patient will respond to a candidate therapeutic intervention aimed at treating clinical symptoms. In a preferred approach, the patient's allele status is rapidly diagnosed using a sensitive PCR assay and a treatment protocol is rendered. The invention also provides a method for forecasting patient outcome and the suitability of the patient for entering a clinical drug trial for the testing of a candidate therapeutic intervention for a disease, condition, or dysfunction as identified herein.

The findings described herein indicate the predictive value of the target allele in identifying patients at risk for a disease or disorder as identified for aspects herein. In addition, because the underlying mechanism influenced by the allele status is not disease-specific, the allele status is suitable for making patient predictions for diseases not affected by the pathway as well.

The following examples, which describe exemplary techniques and experimental results, are provided for the purpose of illustrating the invention, and should not be construed as limiting.

## Example 1

Method for Producing cDNA

In order to identify sequence variances in a gene by laboratory methods it is in some instances useful to produce cDNA(s) from multiple human subjects. (In other instances it may be preferable to study genomic DNA.). Methods for producing cDNA are known to those skilled in the art, as are methods for amplifying and sequencing the cDNA or portions thereof. An example of a useful cDNA

production protocol is provided below. As recognized by those skilled in the art, other specific protocols can also be used.

cDNA Production

** Make sure that all tubes and pipette tips are RNase-free. (Bake them overnight at 100°C in a vacuum oven to make them RNase-free.)

1. Add the following to a RNase-free 0.2 ml micro-amp tube and mix gently:

24 ul water (DEPC treated)
12 ul RNA (1ug/ul)
12 ul random hexamers(50 ng/ul)

2. Heat the mixture to 70°C for ten minutes.

3. Incubate on ice for 1 minute.

4. Add the following:

16 ul 5 X Synthesis Buffer
8 ul 0.1 M DTT
4 ul 10 mM dNTP mix (10 mM each dNTP)
4 ul SuperScript RT II enzyme

Pipette gently to mix.

5. Incubate at 42°C for 50 minutes.

6. Heat to 70°C for ten minutes to kill the enzyme, then place it on ice.

7. Add 160 ul of water to the reaction so that the final volume is 240 ul.

8. Use PCR to check the quality of the cDNA. Use primer pairs that will give a ~800 base pair long piece. See "PCR Optimization" for the PCR protocol.

The following chart shows the reagent amounts for a 20 ul reaction, a 80 ul reaction, and a batch of 39 (which makes enough mix for 36) reactions:

| | 20 ul X 1 tube | 80 ul X 1 tube | 80ul X 39 tubes | |
|---|---|---|---|---|

| | | | | |
|---|---|---|---|---|
| Water | 6 ul | 24 ul | 936 | water |
| RNA | 3 ul | 12 ul | | RNA |
| random hexamers | 3 ul | 12 ul | 468 | random hexamers |
| | | | | |
| synthesis buffer | 4 ul | 16 ul | 624 | synthesis buffer |
| 0.1 M DTT | 2 ul | 8 ul | 312 | 0.1 M DTT |
| 10mM dNTP | 1 ul | 4 ul | 156 | 10mM dNTP |
| SSRT | 1 ul | 4 ul | 156 | SSRT |

**Example 2**

Method for Detecting Variances by Single Strand Conformation Polymorphism (SSCP) Analysis

This example describes the SSCP technique for identification of sequence variances of genes. SSCP is usually paired with a DNA sequencing method, since the SSCP method does not provide the nucleotide identity of variances. One useful sequencing method, for example, is DNA cycle sequencing of $^{32}P$ labeled PCR products using the Femtomole DNA cycle sequencing kit from Promega (WI) and the instructions provided with the kit. Fragments are selected for DNA sequencing based on their behavior in the SSCP assay.

Single strand conformation polymorphism screening is a widely used technique for identifying an discriminating DNA fragments which differ from each other by as little as a single nucleotide. As originally developed by Orita et al. (Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc Natl Acad Sci U S A.* 86(8):2766-70, 1989), the technique was used on genomic DNA, however the same group showed that the technique works very well on PCR amplified DNA as well. In the last 10 years the technique has been used in hundreds of published papers, and modifications of the technique have been described in dozens of papers. The enduring popularity of the technique is due to (1) a high degree of sensitivity to single base differences (>90%) (2) a high degree of selectivity, measured as a low frequency of false positives, and (3) technical ease. SSCP is almost always used together with DNA sequencing because SSCP does not directly provide the sequence basis of differential fragment mobility. The basic steps of the SSCP procdure are described below.

When the intent of SSCP screening is to identify a large number of gene variances it is useful to screen a relatively large number of individuals of different racial, ethnic and/or geographic origins. For example, 32 or 48 or 96 individuals is a convenient number to screen because gel electrophoresis apparatus are available with 96 wells (Applied Biosystems Division of Perkin Elmer Corporation), allowing 3 X 32, 2 X 48 or 96 samples to be loaded per gel.

The 32 (or more) individuals screened should be representative of most of the worlds major populations. For example, an equal distribution of Africans, Europeans and Asians constitutes a reasonable screening set. One useful source of cell lines from different populations is the Coriell Cell Repository (Camden, NJ), which sells EBV immortalized lyphoblastoid cells obtained from several thousand subjects, and includes the racial/ethnic/geographic background of cell line donors in its catalog. Alternatively, a panel of cDNAs can be isolated from any specific target population.

SSCP can be used to analyze cDNAs or genomic DNAs. For many genes cDNA analysis is preferable because for many genes the full genomic sequence of the target gene is not available, however, this circumstance will change over the next few years. To produce cDNA requires RNA. Therefore each cell lines is grown to mass culture and RNA is isolated using an acid/phenol protocol, sold in kit form as Trizol by Life Technologies (Gaithersberg, MD). The unfractionated RNA is used to produce cDNA by the action of a modified Maloney Murine Leukemia Virus Reverse Transcriptase, purchased in kit form from Life Technologies (Superscript II kit). The reverse transcriptase is primed with random hexamer primers to initiate cDNA synthesis along the whole length of the RNAs. This proved useful later in obtaining good PCR products from the 5' ends of some genes. Alternatively, oligodT can be used to prime cDNA synthesis.

Material for SSCP analysis can be prepared by PCR amplification of the cDNA in the presence of one $\alpha$ $^{32}$P labeled dNTP (usually $\alpha$ $^{32}$P dCTP). Usually the concentration of nonradioactive dCTP is dropped from 200 uM (the standard concentration for each of the four dNTPs) to about 100 uM, and $^{32}$P dCTP is added to a concentration of about 0.1-0.3 uM. This involves adding a 0.3- 1 ul (3-10 uCi) of $^{32}$P cCTP to a 10 ul PCR reaction. Radioactive nucleotides can be purchased from DuPont/New England Nuclear.

The customary practice is to amplify about 200 base pair PCR products for SSCP, however, an alternative approach is to amplify about 0.8-1.4 kb fragments and then use several cocktails of restriction endonucleases to digest those into smaller fragments of about 0.1-0.4kb, aiming to have as many fragments as possible between .15 and .3 kb. The digestion strategy has the advantage that less PCR is required, reducing both time and costs. Also, several different restriction enzyme digests can be performed on each set of samples (for example 96 cDNAs), and then each of the digests can be run separately on SSCP gels. This redundant method (where each nucleotide is surveyed in three different fragments) reduces both the false negative and false positive rates. For example: a site of variance might lie within 2 bases of the end of a fragment in one digest, and as a result not affect the

conformation of that strand; the same variance, in a second or third digest, would likely lie in a location more prone to affect strand folding, and therefore be detected by SSCP.

After digestion, the radiolabelled PCR products are diluted 1:5 by adding formamide load buffer (80% formamide, 1X SSCP gel buffer) and then denatured by heating to 90%C for 10 minutes, and then allowed to renature by quickly chilling on ice. This procedure (both the dilution and the quick chilling) promotes intra- (rather than inter-) strand association and secondary structure formation. The secondary structure of the single strands influences their mobility on nondenaturing gels, presumably by influencing the number of collisions between the molecule and the gel matrix (i.e., gel sieving). Even single base differences consistently produce changes in intrastrand folding sufficient to register as mobility differences on SSCP.

The single strands were then resolved on two gels, one a 5.5% acrylamide, 0.5X TBE gel, the other an 8% acrylamide, 10% glycerol, 1X TTE gel. (Other gel recipes are known to those skilled in the art.) The use of two gels provides a greater opportunity to recognize mobility differences. Both glycerol and acrylamide concentration have been shown to influence SSCP performance. By routinely analyzing three different digests under two gel conditions (effectively 6 conditions), and by looking at both strands under all 6 conditions, one can achieve a 12-fold sampling of each base pair of cDNA. However, if the goal is to rapidly survey many genes or cDNAs then a less redundant procedure would be optimal.

**Example 3**

Method for Detecting Variances by T4 endonuclease VII (T4E7) mismatch cleavage method

The enzyme T4 endonuclease VII is derived from the bacteriophage T4. T4 endonuclease VII is used by the bacteriophage to cleave branched DNA intermediates which form during replication so the DNA can be processed and packaged. T4 endonuclease can also recognize and cleave heteroduplex DNA containing single base mismatches as well as deletions and insertions. This activity of the T4 endonuclease VII enzyme can be exploited to detect sequence variances present in the general population.

The following are the major steps involved in identifying sequence variations in a candidate gene by T4 endonuclease VII mismatch cleavage:

1. Amplification by the polymerase chain reaction (PCR) of 400-600 bp regions of the candidate gene from a panel of DNA samples The DNA samples can either be cDNA or genomic DNA and will represent some cross section of the world population.
2. Mixing of a fluorescently labeled probe DNA with the sample DNA. Heating and cooling the mixtures causing heteroduplex formation between the probe DNA and the sample DNA.
3. Addition of T4 endonuclease VII to the heteroduplex DNA samples. T4 endonuclease will recognize and cleave at sequence variance mismatches formed in the heteroduplex DNA.
4. Electrophoresis of the cleaved fragments on an ABI sequencer to determine the site of cleavage.
5. Sequencing of a subset of PCR fragments identified by T4 endonuclease VI to contain variances to establish the specific base variation at that location.

A more detailed description of the procedure is as follows:

A candidate gene sequence is downloaded from an appropriate database. Primers for PCR amplification are designed which will result in the target sequence being divided into amplification products of between 400 and 600 bp. There will be a minimum of a 50 bp of overlap not including the primer sequences between the 5' and 3' ends of adjacent fragments to ensure the detection of variances which are located close to one of the primers.

Optimal PCR conditions for each of the primer pairs is determined experimentally. Parameters including but not limited to annealing temperature, pH, $MgCl_2$ concentration, and KCl concentration will be varied until conditions for optimal PCR amplification are established. The PCR conditions derived for each

primer pair is then used to amplify a panel of DNA samples (cDNA or genomic DNA) which is chosen to best represent the various ethnic backgrounds of the world population or some designated subset of that population.

One of the DNA samples is chosen to be used as a probe. The same PCR conditions used to amplify the panel are used to amplify the probe DNA. However, a flourescently labeled nucleotide is included in the deoxy-nucleotide mix so that a percentage of the incorporated nucleotides will be fluorescently labeled.

The labeled probe is mixed with the corresponding PCR products from each of the DNA samples and then heated and cooled rapidly. This allows the formation of heteroduplexes between the probe and the PCR fragments from each of the DNA samples. T4 endonuclease VII is added directly to these reactions and allowed to incubate for 30 min. at 37 C. 10 ul of the Formamide loading buffer is added directly to each of the samples and then denatured by heating and cooling. A portion of each of these samples is electrophoresed on an ABI 377 sequencer. If there is a sequence variance between the probe DNA and the sample DNA a mismatch will be present in the heteroduplex fragment formed. The enzyme T4 endonuclease VII will recognize the mismatch and cleave at the site of the mismatch. This will result in the appearance of two peaks corresponding to the two cleavage products when run on the ABI 377 sequencer.

Fragments identified as containing sequencing variances are subsequently sequenced using conventional methods to establish the exact location and sequence variance.

**Example 4**

Method for Detecting Variances by DNA sequencing.

Sequencing by the Sanger dideoxy method or the Maxim Gilbert chemical cleavage method is widely used to determine the nucleotide sequence of genes. Presently, a worldwide effort is being put forward to sequence the entire human genome. The Human Genome Project as it is called has already resulted in the identification and sequencing of many new human genes. Sequencing can not only be used to identify new genes, but can also be used to identify variations between individuals in the sequence of those genes.

The following are the major steps involved in identifying sequence variations in a candidate gene by sequencing:

1. Amplification by the polymerase chain reaction (PCR) of 400-700 bp regions of the candidate gene from a panel of DNA samples The DNA samples can either be cDNA or genomic DNA and will represent some cross section of the world population.

2. Sequencing of the resulting PCR fragments using the Sanger dideoxy method. Sequencing reactions are performed using flourescently labeled dideoxy terminators and fragments are separated by electrophoresis on an ABI 377 sequencer or its equivalent.

3. Analysis of the resulting data from the ABI 377 sequencer using software programs designed to identify sequence variations between the different samples analyzed.

A more detailed description of the procedure is as follows:

A candidate gene sequence is downloaded from an appropriate database. Primers for PCR amplification are designed which will result in the target sequence being divided into amplification products of between 400 and 700 bp. There will be a minimum of a 50 bp of overlap not including the primer sequences between the 5' and 3' ends of adjacent fragments to ensure the detection of variances which are located close to one of the primers.

Optimal PCR conditions for each of the primer pairs is determined experimentally. Parameters including but not limited to annealing temperature, pH, $MgCl_2$ concentration, and KCl concentration will be varied until conditions for optimal PCR amplification are established. The PCR conditions derived for each primer pair is then used to amplify a panel of DNA samples (cDNA or genomic

DNA) which is chosen to best represent the various ethnic backgrounds of the world population or some designated subset of that population.

PCR reactions are purified using the QIAquick 8 PCR purification kit (Qiagen cat# 28142) to remove nucleotides, proteins and buffers. The PCR reactions are mixed with 5 volumes of Buffer PB and applied to the wells of the QIAquick strips. The liquid is pulled through the strips by applying a vacuum. The wells are then washed two times with 1 ml of buffer PE and allowed to dry for 5 minutes under vacuum. The PCR products are eluted from the strips using 60 ul of elution buffer.

The purified PCR fragments are sequenced in both directions using the Perkin Elmer ABI Prism™ Big Dye™ terminator Cycle Sequencing Ready Reaction Kit (Cat# 4303150). The following sequencing reaction is set up: 8.0 ul Terminator Ready Reaction Mix, 6.0 ul of purified PCR fragment, 20 picomoles of primer, deionized water to 20 ul. The reactions are run through the following cycles 25 times: 96°C for 10 second, annealing temperature for that particular PCR product for 5 seconds, 60°C for 4 minutes.

The above sequencing reactions are ethanol precipitated directly in the PCR plate, washed with 70% ethanol, and brought up in a volume of 6 ul of formamide dye. The reactions are heated to 90°C for 2 minutes and then quickly cooled to 4°C. 1 ul of each sequencing reaction is then loaded and run on an ABI 377 sequencer.

The output for the ABI sequencer appears as a series of peaks where each of the different nucleotides. A, C, G, and T appear as a different color. The nucleotide at each position in the sequence is determined by the most prominent peak at each location. Comparison of each of the sequencing outputs for each sample can be examined using software programs to determine the presence of a variance in the sequence. One example of heterozygote detection using sequencing with dye labeled terminators is described by Kwok *et. al.* (Kwok, P.-Y.; Carlson, C.; Yager, T.D., Ankener, W.,and D. A. Nickerson, *Genomics* 23, 138-144, 1994). The software compares each of the normalized peaks between all the samples base by base and looks for a 40% decrease in peak height and the concomitant appearance of a new peak underneath. Possible variances flagged by the software are further analyzed visually to confirm their validity.

**Example 5**

Hardy-Weinberg equilibrium

Evolution is the process of change and diversification of organisms through time, and evolutionary change affects morphology, physiology and reproduction of organisms, including humans. These evolutionary changes are the result of changes in the underlying genetic or hereditary material. Evolutionary changes in a group of interbreeding individuals or Mendelian population, or simply populations, are described in terms of changes in the frequency of genotypes and their constituent alleles. Genotype frequencies for any given generation is the result of the mating among members (genotypes) of their previous generation. Thus, the expected proportion of genotypes from a random union of individuals in a given population is essential for describing the total genetic variation for a population of any species. For example, the expected number of genotypes that could form from the random union of two alleles, A and a, of a gene are AA, Aa and aa. The expected frequency of genotypes in a large, random mating population was discovered to remain constant from generation to generation; or achieve Hardy-Weinberg equilibrium, named after its discoverers. The expected genotypic frequencies of alleles A and a (AA, 2Aa, aa) are conventionally described in terms of $p^2 + 2pq + q^2$ in which p and q are the allele frequencies of A and a. In this equation ($p^2 + 2pq + q^2 = 1$), p is defined as the frequency of one allele and q as the frequency of another allele for a trait controlled by a pair of alleles (A and a). In other words, p equals all of the alleles in individuals who are homozygous dominant (AA) and half of the alleles in individuals who are heterozygous (Aa) for this trait. In mathematical terms, this is

$$p = AA + \tfrac{1}{2}Aa$$

Likewise, q equals the other half of the alleles for the trait in the population, or

$$q = aa + \tfrac{1}{2}Aa$$

Because there are only two alleles in this case, the frequency of one plus the frequency of the other must equal 100%, which is to say

$$p + q = 1$$

Alternatively,

$$p = 1 - q \text{ OR } q = 1 - p$$

All possible combinations of two alleles can be expressed as:

$$(p + q)^2 = 1$$

or more simply,

$$p^2 + 2pq + q^2 = 1$$

In this equation, if p is assumed to be dominant, then $p^2$ is the frequency of homozygous dominant (AA) individuals in a population, 2pq is the frequency of

heterozygous (Aa) individuals, and $q^2$ is the frequency of homozygous recessive (aa) individuals.

From observations of phenotypes, it is usually only possible to know the frequency of homozygous dominant or recessive individuals, because both dominant and recessives will express the distinguishable traits. However, the Hardy-Weinberg equation allows us to determine the expected frequencies of all the genotypes, if only p or q is known. Knowing p and q, it is a simple matter to plug these values into the Hardy-Weinberg equation ($p^2 + 2pq + q^2 = 1$). This then provides the frequencies of all three genotypes for the selected trait within the population. This illustration shows Hardy-Weinberg frequency distributions for the genotypes AA, Aa, and aa at all values for frequencies of the alleles, p and q. It should be noted that the proportion of heterozygotes increases as the values of p and q approach 0.5.

*Linkage disequilibirum*

Linkage is the tendency of genes or DNA sequences (e.g. SNPs) to be inherited together as a consequence of their physical proximity on a single chromosome. The closer together the markers are, the lower the probability that they will be separated during DNA crossing over, and hence the greater the probability that they will be inherited together. Suppose a mutational event introduces a "new" allele in the close proximity of a gene or an allele. The new allele will tend to be inherited together with the alleles present on the "ancestral," chromosome or haplotype. However, the resulting association, called linkage disequilibrium, will decline over time due to recombination. Linkage disequilibrium has been used to map disease genes. In general, both allele and haplotype frequencies differ among populations. Linkage disequilibrium is varied among the populations, being absent in some and highly significant in others.

*Quantification of the relative risk of observable outcomes of a Pharmacogenetics Trial*

Let PlaR be the placebo response rate (0% ( PlaR ( 100%) and TntR be the treatment response rate (0% ( TntR ( 100%) of a classical clinical trial. ObsRR is defined as the relative risk between TntR and PlaR:

$$\text{ObsRR} = \text{TntR} / \text{PlaR}.$$

Suppose that in the treatment group there is a polymorphism in relation to drug metabolism such as the treatment response rate is different for each genotypic subgroup of patients. Let q be the allele a frequency of a recessive biallelic locus (e.g. SNP) and $p = 1 - q$ the allele A frequency. Following Hardy-Weinberg

equilibrium, the relative frequency of homozygous and heterozygous patients are as follow:

AA: p2 Aa: 2pq aa: q2

with

$$(p2+2pq+q2) = 1.$$

Let's define AAR, AaR, aaR as respectively the response rates of the AA, Aa and aa patients. We have the following relationship:

$$TntR = AAR*p2 + AaR*2pq + aaR*q2.$$

Suppose that the aa genotypic group of patients has the lowest response rate, i.e. a response rate equal to the placebo response rate (which means that the polymorphism has no impact on natural disease evolution but only on drug action) and let's define ExpRR as the relative risk between AAR and aaR, as

$$ExpRR = AAR / aaR.$$

From the previous equations, we have the following relationships:

ObsRR ( ExpRR ( 1/PlaR

$$TntR / PlaR = (AAR*p2 + AaR*2pq + aaR*q2) / PlaR$$

The maximum of the expected relative risk, max(ExpRR), corresponding to the case of heterozygous patients having the same response rate as the placebo rate, is such that:

$$ObsRR = ExpRR*p2 + 2pq + q2 \quad \Leftrightarrow \quad ExpRR = (ObsRR - 2pq - q2) / p2$$

The minimum of the expected relative risk, min(ExpRR), corresponding to the case of heterozygous patients having the same response rate as the homozygous non-affected patients, is such that:

$$ObsRR = ExpRR*(p2 + 2pq) + q2 \quad \Leftrightarrow \quad ExpRR = (ObsRR - q2) / (p2 + 2pq)$$

For example, if q = 0.4, PlaR = 40% and ObsRR = 1.5 (i.e. TntR = 60%), then 1.6 ( ExpRR ( 2.4. This means that the best treatment response rate we can expect in a genotypic subgroup of patients in these conditions would be 95.6% instead of 60%.

This can also be expressed in terms of maximum potential gain between the observed difference in response rates (TntR – PlaR) without any pharmacogenetic hypothesis and the maximum expected difference in response rates (max(ExpRR)*PlaR – TntR) with a strong pharmacogenetic hypothesis:

$$(max(ExpRR)*PlaR - TntR) = [(ObsRR - 2pq - q2) / p2] * PlaR - TntR$$

$$\Leftrightarrow \quad (max(ExpRR)*PlaR - TntR) = [TntR - PlaR*(2pq + q2) - TntR*p2]/p2$$

$$\Leftrightarrow \quad (max(ExpRR)*PlaR - TntR) = [TntR*(1- p2) - PlaR*(2pq + q2)]/p2$$

$$\Leftrightarrow \quad (max(ExpRR)*PlaR - TntR) = [(1 - p2) / p2] * (TntR - PlaR)$$

that is for the previous example,

$$(95.6\% - 60\%) = [(1 - 0.62)/0.62]^* (60\% - 40\%) = 35.6\%$$

Suppose that, instead of one SNP, we have p loci of SNPs for one gene. This means that we have 2p possible haplotypes for this gene and (2p)(2p-1)/2 possible genotypes. And with 2 genes with p1 and p2 SNP loci, we have [(2p1)(2p1-1)/2]*[(2p2)(2p2-1)/2] possibilities; and so on. Examining haplotypes instead of combinations of SNPs is especially useful when there is linkage disequilibrium enough to reduce the number of combinations to test, but not complete since in this latest case one SNP would be sufficient. Yet the problem of frequency above still remains with haplotypes instead of SNPs since the frequency of a haplotype cannot be higher than the highest SNP frequency involved.

*Statistical Methods to be used in Objective Analyses*

The statistical significance of the differences between variance frequencies can be assessed by a Pearson chi-squared test of homogeneity of proportions with *n-1* degrees of freedom. Then, in order to determine which variance(s) is(are) responsible for an eventual significance, we can consider each variance individually against the rest, up to *n* comparisons, each based on a 2x2 table. This should result in chi-squared tests that are individually valid, but taking the most significant of these tests is a form of multiple testing. A Bonferroni's adjustment for multiple testing will thus be made to the P-values, such as $p^*=1-(1-p)^n$.

The statistical significance of the difference between genotype frequencies associated to every variance can be assessed by a Pearson chi-squared test of homogeneity of proportions with 2 degrees of freedom, using the same Bonferroni's adjustment as above.

Testing for unequal haplotype frequencies between cases and controls can be considered in the same framework as testing for unequal variance frequencies since a single variance can be considered as a haplotype of a single locus. The relevant likelihood ratio test compares a model where two seqarate sets of haplotype frequencies apply to the cases and controls, to one where the entire sample is characterized by a single common set of haplotype frequencies. This can be performed by repeated use of a computer program (Terwilliger and Ott, 1994, Handbook of Human Linkage Analysis, Baltimore, John Hopkins University Press) to successively obtain the log-likelihood corresponding to the set of haplotpe frequency estimates on the cases ($\ln L_{case}$), on the controls ($\ln L_{control}$), and on the overall ($\ln L_{combined}$). The test statistic $2((\ln L_{case})+ (\ln L_{control})- (\ln L_{combined}))$ is then chi-squared with *r-1* degrees of freedom (where *r* is the number of haplotypes).

To test for potentially confounding effects or effect-modifiers, such as sex, age, etc., logistic regression can be used with case-control status as the outcome variable, and genotypes and covariates (plus possible interactions) as predictor variables.

5

**Example 6**

Exemplary Pharmacogenetic Analysis Steps

In accordance with the discussion of distribution frequencies for variances, alleles, and haplotypes, variance detection, and correlation of variances or haplotypes with treatment response variability, the points below list major items which will typically be performed in an analysis of the pharmacogenetic determination of the effects of variances in the treatment of a disease and the selection/optimization of treatment.

1) List candidate gene/genes for a known genetic disease, and assign them to the respective metabolic pathways.

2) Determine their alleles, observed and expected frequencies, and their relative distributions among various ethnic groups, gender, both in the control and in the study (case) groups.

3) Measure the relevant clinical/phenotypic (biochemical / physiological) variables of the disease.

4) If the causal variance/allele in the candidate gene is unknown, then determine linkage disequilibria among variances of the candidate gene(s).

5) Divide the regions of the candidate genes into regions of high linkage disequilibrium and low disequilibrium.

6) Develop haplotypes among variances that show strong linkage disequilibrium using the computation methods.

7) Determine the presence of rare haplotypes experimentally. Confirm if the computationally determined rare haplotypes agree with the experimentally determined haplotypes.

8) If there is a disagreement between the experimentally determined haplotypes and the computationally derived haplotypes, drop the computationally derived rare haplotypes, construct cladograms from these haplotypes using the Templeton (1987) algorithm.

9) Note regions of high recombination. Divide regions of high recombination further to see patterns of linkage disequilibria.

10) Establish association between cladograms and clinical variables using the nested analysis of variance as presented by Templeton (1995), and assign causal variance to a specific haplotype.

11) For variances in the regions of high recombination, use permutation tests for establishing associations between variances and the phenotypic variables.

12) If two or more genes are found to affect a clinical variable determine the relative contribution of each of the genes or variances in relation to the clinical variable, using step-wise regression or discriminant function or principal component analysis.

13) Determine the relative magnitudes of the effects of any of the two variances on the clinical variable due to their genetic (additive, dominant or epistasis) interaction.

14) Using the frequency of an allele or haplotypes, as well as biochemical/clinical variables determined in the *in vitro* or *in vivo* studies, determine the effect of that gene or allele on the expression of the clinical variable, according to the measured genotype approach of Boerwinkle et al (Ann. Hum. Genet 1986).

15) Stratify ethnic/ clinical populations based on the presence or absence of a given allele or a haplotype.

16) Optimize drug dosages based on the frequency of alleles and haplotypes as well as their effects using the measured genotype approach as a guide.

**Example 7**

Exemplary Pharmacogenetic Analysis Steps - biological function analysis

In many cases when a gene which may affect drug action is found to exhibit variances in the gene, RNA, or protein sequence, it is preferable to perform biological experiments to determine the biological impact of the variances on the structure and function of the gene or its expressed product and on drug action. Such experiments may be performed *in vitro* or *in vivo* using methods known in the art.

The points below list major items which may typically be performed in an analysis of the effects of variances in the treatment of a disease and the selection/optimization of treatment using biological studies to determine the structure and function of variant forms of a gene or its expressed product..

1) List candidate gene/genes for a known genetic disease, and assign them to the respective metabolic pathways.

2) Identify variances in the gene sequence, the expressed mRNA sequence or expressed protein sequence.

3) Match the position of variances to regions of the gene, mRNA, or protein with known biological functions. For example, specific sequences in the promotor of a gene are known to be responsible for determining the level of expression of the gene; specific sequences in the mRNA are known to be involved in the processing of nuclear mRNA into cytoplasmic mRNA including splicing and polyadenylation; and certain sequences in proteins are known to direct the trafficking of proteins to specific locations within a cell and to constitute active sites of biological functions including the binding of proteins to other biological consituents or catalytic functions. Variances in sites such as these, and others known in the art, are candidates for biological effects on drug action.

4) Model the effect of the variance on mRNA or protein structure. Computational methods for predicting the structure of mRNA are known and can be used to assess whether a specific variance is likely to cause a substantial change in the structure of mRNA. Computational methods can also be used to predict the structure of peptide sequences enabling predictions to be made concerning the potential impact of the variance on protein function. Most useful are structures of proteins determined by X-ray diffraction, NMR or other methods known in the art which provide the atomic structure of the protein. Computational methods can be used to consider the effect of changing an amino acid within such a structure to determine whether such a change would disrupt the structure and/or funciton of the protein. Those skilled in the art will recognize that this analysis can be performed on crystal structures of the protein known to have a variance as well as homologous proteins expressed from different loci in the human genome, or homologous proteins from other species, or non-homologous but analogous proteins with similar functions from humans or other species.

5) Produce the gene, mRNA or protein in amounts sufficient to experimentally characterize the structure and function of the gene, mRNA or protein. It will be apparent to those skilled in the art that by comparing the activity of two genes or their products which differ by a single variance, the effect of the variance can be determined. Methods for producing genes or gene products which differ by one or more bases for the purpose of experimental analysis are known in the art.

6) Experimental methods known in the art can be used to determine whether a specific variance alters the transcription of a gene and translation into a gene product. This involves producing amounts of the gene by molecular cloning sufficient for in vitro or in vivo studies. Methods for producing genes and gene products are known in the art and include cloning of segments of genetic material in prokaryotes or eukarotic hosts, run off transcription and cell-free translation assays that can be performed in cell free extracts, transfection of DNA into cultured cells, introduction of genes into live animals or embryos by direct injection or using vehicles for gene delivery including transfection mixtures or viral vectors.

7) Experimental methods known in the art can be used to determine whether a specific variance alters the ability of a gene to be transcribed into RNA. For example, run off transcription assays can be performed in vitro or expression can be characterized in transfected cells or transgenic animals.

8) Experimental methods known in the art can be used to determine whether a specific variance alters the processing, stability, or translation of RNA into protein. For example, reticulocyte lysate assays can be used to study the production of protein in cell free systems, transfection assays can be designed to study the production of protein in cultured cells, and the production of gene products can be measured in transgenic animals.

9) Experimental methods known in the art can be used to determine whether a specific variant alters the activity of an expressed protein product. For example, protein can be producted by reticulocyte lystae systems or by introducing the gene into prokaryotic organisms such as bacteria or lowre eukaryotic organisms such as yeast or fungus), or by introducing the gene into cultured cells or transgenic animals. Protein produced in such systems can be extracted or purified and subjected to bioassays known to those in the art as measures of the nction of that particular protein. Bioassays may involve, but are not limited to, binding, inhibiton, or catalytic functions.

10) Those skilled in the art will recognize that it is sometimes preferred to perform the above experiments in the presence of a specific drug to determine whether the drug has differential effects on the activity being measured. Alternatively, studies may be performed in the presence of an analogue or metabolite of the drug.

11) Using methods described above, specific variances which alter the biological function of a gene or its gene product that could have an impact on drug action can be identified. Such variances are then studied in clinical trial populations to determine whether the presence or absence of a specific variance correlates with observed clinical outcomes such as efficacy or toxicity.

12) It will be further recognized that there may be more than one variance within a gene that is capable of altering the biological function of the gene or gene product. These variances may exhibit similar, synergistic effects, or may have opposite effects on gene function. In such cases, it is necessary to consider the haplotype of the gene, namely the combination of variances that are present within a single allele, to assess the composite function of the gene or gene product.

13) Perform clinical trials with stratification of patients based on presence or absence of a given variance, allele or haplotype of a gene. Establish associations between observed drug responses such as toxicity, efficacy, drug response, or dose toleration and the presence or absense of a specific variance, allele, or haplotype.

14) Optimize drug dosage or drug usage based on the presence of the variant.

**Example 8**

Stratification of patients by genotype in prospective clinical trials.

In a prospective clinical trial, patients will be stratified by genotype to determine whether the observed outcomes are different in patients having different genotypes. A critical issue is the design of such trials to assure that a sufficient number of patients are studied to observe genetic effects.

The number of patients required to achieve statistical significance in a conventional clinical trial is calculated from:

1:1 $N=2(z_\alpha+z_{2\beta})^2 / (\delta/\sigma)^2$ (two tailed test)

From this equation it may be inferred that the size of a genetically defined subgroup $N_i$ required to achieve statistical significance for an observed outcome associated with variance or haplotype "i" can be calculated as:

1.2 $N_i=2(z_\alpha+z_{2\beta})^2 / (\delta_i/\sigma_i)^2$

If $P_i$ is the prevalence of the genotype "i"in the population, the total number of patients that need to be incorporated in a clinical trial $N_g$ to identify a population with haplotype "i" of size $N_i$ is given by:

1.3 $N_g=N_i/P_i$

It should be noted that $N_g$ describes the total number of patients that need to be genotyped in order to identify a subset of $N_i$ patients with genotype "i".

If genotyping is used as means for statistical stratification of patients, $N_g$ represents the number of patients that would need to be enrolled in a trial to achieve statistical significance for subgroup "i". If genotyping is used as a means for inclusion, it represents the number of patients that need to screened to identify a population of $N_i$ individuals for an appropriately powered clinical trial. Thus, $N_g$ is a critical determinant of the scope of the clinical trial as well as $N_i$.

A clinical trial can also be designed to test associations for multiple genetic subgroups "j" defined by a single allele in which case:

1.4 $N_g = \max ( N_{gi} )$ for i=1...j

If more than one subgroup is tested, but there is no overlap in the patients contained within the subgroups, these can be considered to be independent hypotheses and no multiple testing correction should be required. If consideration of more than one subgroup constitutes multiple testing, or if individual patients are included in multiple subgroups, then statistical corrections may required in the values of $z_\alpha$ or $z_{2\beta}$ which would increase the number of patients required.

It should be emphasized that a clinical trial of this nature may not provide statistically significant data concerning associations with any genotype other than "i". The total number of patients that would be required in a clinical trial to test

more than one genetically defined subgroup would be determined by the maximum value of $N_g$ for any single subgroup.

The power of pharmacogenomics to improve the efficiency of clinical trials arises from the fact it is possible to have $N_g<N$. The goal of pharmacogenomic analysis is to identify a genetically define subgroup in which the magnitude of the clinical response is greater and the variability in response is reduced. These observations correspond to an increase in the magnitude of the (mean) observed response $\delta$ or a decrease the degree of variability $\sigma$. Since the value of $N_i$ calculated in equation 1.2 decreases non-linearly as the square of these changes, the total number of patients $N_g$ can also decrease non-linearly, resulting in a clinical trial that requires fewer patients to achieve statistical significance. If $\delta_i$and $\sigma_i$ are not different than $\delta$ and $\sigma$, then $N_g$ is greater than N as given by $N_g=N_i/P_i$. Values of $\delta_i$and $\sigma_i$ that give $N_g<N$ can be calculated:

1.5 $N_g<N$ if: $P_i> [(\delta/\sigma)^2]/[(\delta_i/\sigma_i)^2]$

It is apparent from this analysis that $N_g$ is not uniformly less than N, even with modest improvements in the values for $\delta_i$and $\sigma_i$.

As with a conventional clinical trial, the incorporation of an appropriate control group in the study design is critical for achieving success. In the case of a prospective clinical trial, the control group commonly is selected on the basis of the same inclusion criteria as the treatment group, but is treated with placebo or a standard therapeutic regimen rather than the investigational drug. In the case of a study with subgroups that are defined by haplotype, the ideal control group for a treatment subgroup with hapotype "i" is a placebo-treated subgroup with haplotype "i". This is often a critical control, since haplotypes which may be associated with the response to treatment may also affect the natural course of the disease.

A critical issue in considering control groups is that $\sigma$ for the control group placebo treated population with haplotype "i" may not be equivalent to that of the control population. If so, 1.5 may overestimate the benefits of any reduction in $\sigma_i$ in the treatment response group if there is not also a reduction in $\sigma_i$ in the control group.

If $\sigma$ of the treatment and control groups are not equivalent, $\delta$ would be still calculated as the difference in the response of the two groups, but $\sigma$ would be

different in the two groups with values of $\sigma_0$ or $\sigma_1$ respectively. In this case, the number of patients in the genetically defined subgroup $N_i$ would be defined by:

2.1 $N_i=(\sigma Z_\alpha + \sigma_i Z_\beta)^2/\delta^2$

The total number of patients that would need to be enrolled in such a trial would be the maximium of

2.2 N or N/Pi

It will be apparent that such an analysis remains sensitive to increases in $\delta$, but is less sensitive to changes in $\sigma$ which are not also reflected in the control group.

Certain analysis may be performed by comparing individuals with one haplotype against the entire normal population. Such an analysis may be used to establish the selectivity of the response associated with a specific haplotype. For example, it may be desirable to establish that the response or toxicity observed in a specific subgroup is greater than that associated observed with the entire population. It may also be of interest to compare the response to treatment between two different subgroups. If $\sigma$ differs between the groups, then the estimate of the number of patients that need to be enrolled in the trial must be calculated using equations 2.1 with N being the maximum of $N_i/P_i$ for the different subgroups.

Another issue in controls is the relative size of the treatment and control groups. In a prospectively designed clinical trial which selectively incorporates patients with haplotype "i"the number of patients in the control and treatment group will be essentially equivalent. If the control group is different, or if haplotypes are used for stratification but not inclusion, statistical corrections may need to be made for having populations of different size.

**Example 9**

Stratification of patients by phenotype.

The identification of genetic associations in Phase II or retrospective studies can be performed by stratifying patients by phenotype and analyzing the distribution of genotypes/haplotypes in the separate populations. A particularly important aspect of this analysis is that any gene may have only a partial effect on the observed outcome, meaning that there will be an association value (A) corresponding to the

fraction of patients in a phenotypically-defined subgroup who exhibit that phenotype due to a specific genotype/phenotype.

It will be recognized to those skilled in the art that the fraction of individuals who exhibit a phenotype due to any specific allele will be less than 1 (i.e. A<1). This is true for several reasons. The observed phenotype may occur by random chance. The observed phenotype may be associated with environmental influences, or the observed phenotype may be due to different genetic effects in differen tindividuals.Furthermore, the onstruction of haplotypes and analysis of recombination may not group all alleles with pheontypically-significant variances within a single haplotype or haplotype cluster. In this case, causative variances at a single locus may be associated with more than one haplotype or haplotype cluster and the association constant A for the locus would be $A=A_1+A_2+...+A_n<1$. It is likely that many phenotypes will be associated with multiple alleles at a given locus, and it is particularly important that statistical methods be sufficiently robust to identify association with a locus even if $A_i$ is reduced by the presence of several causative alleles.

Statistical methods can be used to identify genetic effects on an observed outcome in patient groups stratified by phenotype, eg the presence or absence of the observed response. One such method entails determining the allele frequencies in two populations of patients stratified by an observed clinical outcome, for example efficacy or toxicity and performing a maximum likelihood analysis for the association between a given gene and the observed phenotype based on the allele frequencies and a range of values for A (the association constant between a specific allele and the observed outcome used to stratify patients).

This analysis is performed by comparing the observed gene frequencies in a patient population with an observed outcome to gene frequencies in a table in which the predicted frequencies of different alleles of the gene assuming different values of the association constant A for that allele. This table of predicted gene frequencies can be constructed by those skilled in the art based on the frequency of any specific allele in the normal population, the predicted inheritance of the effect (e.g. dominant or recessive) and the fraction of a subgroup with a specific outcome who would have that allele based on the association constant A.

For example, if a specific outcome was only observed in the presence of a specific allele of a gene, the expected frequency would be 1. If a specific outcome was never observed in the presence of a specific allele of a gene, the expected fequency would be 0. If there was no association between the allele and the observed outcome, the frequency of that allele among individuals with an observed outcome would be the same as in the general population. A statistical analysis can

be performed to compare the observed allele frequencies with the predicted allele frequencies and determine the best fit or maximum likeihood of the association. For example, a chi square analysis will determine whether the observed outcome is statistically similar to predicted outcomes calculated for different modes of inheritance and different potential values of A. P values can then be calculated to determine the likelihood that any specific association is statistically significant. A curve can be calculated based on different values of A, and the maximal likelihood of an association determined from the peak of such a curve. Methods for chi square analysis are known to those in the art.

A multidimensional analysis can also be performed to determine whether an observed outcome is associated with more than one allele at a specific genetic locus. An example of this analysis considering the potential effects of two different alleles of a single gene is shown. It will be apparent to those skilled in the art that this analysis can be extended to n dimensions using computer programs.

This analysis can be used to determine the maximum likelihood that one or more alleles at a given locus are associated with a specific clinical outcome.

It will be apparent to those skilled in the art that critical issues in this analysis include the fidelity of the phenotypic association and identification of a control group. In particular, it may be useful to perform an identical analysis in patients receiving a placebo to eliminate other forms of bias which may contribute to statistical errors.

**Example 10**

Amyotrophic Lateral Sclerosis

I. *Description of Amyotrophic Lateral Sclerosis*

Amyotrophic Lateral Sclerosis (ALS) is a degenerative neurological disease that primarily involves the motor neuron system. The disease is characterized by muscular atrophy, progressive weakness, fasciculations, spasticity, disarthria, dysphagia, and repiratory compromise. Sensory, cognitive, oculomotor, and autonomic functions are spared. There are approximately 30,000 individuals with ALS in the U.S. with an estimated annual cost of $300 million dollars. The majority of cases are sporadic and of unknown etiology, however approximately 5%-10% of ALS cases are inherited as an autosomal dominant trait (familial ALS). Superoxide dismutase 1 (SOD1) gene mutations are responsible for about 20% of familial ALS cases.

II. *Current therapies for ALS*

There are no compounds that halt or prevent the progressive neurodegeneration of ALS. Riluzole (RILUTEK®) , a benzothiazole derivative, is approved for treatment of ALS based on data that it slows disease progression and modestly increases survival time and ventilator-free time.. Riluzole's mechanism of action is not completely understood, however pharmacological properties include: 1) an inhibitory effect on glutamate release, 2) inactivation of voltage dependent sodium channels, 3) downmodulation of signalling via excitatory amino acid receptors, particularly glutamate receptors. Unfortunately riluzole, which was introduced in 1996, produces a benefit in only a fraction of patients, and the effect is modest. For example, despite the increase in longevity there is no consistent increase in muscular strenght or pulmonary function. Thus patients do not experience significant relief from symptoms. Patients and care givers quickly understood these limitations, and consequently the use of the drug has been limited. A 1997 study, conducted during the first 8 months after commercialization of riluzole, found that only 37% of patients (17 of 46) eligible for riluzole were interested in trying the drug. The most common reason given for not wanting to try riluzole was insufficient benefit.

III. *Limitations of Current Therapies for ALS*

As noted above, despite therapy with riluzole, in most ALS patients the disease progresses to debilitating and ultimately life-threatening symptoms. However, since there are no therapeutic alternatives, riluzole is frequently administered despite the modest efficacy. This practice increases the cost of ALS care significantly. In addition to unimpressive efficacy, riluzole therapy has been associated with elevation of serum ALT levels. Thus patients on riluzole should be monitored bimonthly for elevated liver enzymes, at significant cost. Other side effects, which occur infrequently, include neutropenia, asthenia, nausea, dizziness, decreased lung function, diarrhea, abdominal pain, pneumonia, vomiting, vertigo, paresthesia, anorexia, and somnolence. Attending to these iatrogenic effects further increases the costs associated with rizulin therapy.

IV. *Potential Impact of Genotyping on Drug Development for ALS*

There is already a well established genetic cause of some familial ALS cases: mutation of the SOD-1 gene. It is likely that genetic factors play a role in the pathogenesis of sporadic ALS and non-SOD1 linked familial ALS. Strong candidate genes include, for example, other scavengers of superoxide, the entire glutamate signal transduction pathway, calcium channels and genes involved in the production and degradation of neurofilaments. Stratification of clinical trial patients

by allelic variation in these or other candidate genes may reveal differences in response rate, duration or quality of response, or adverse events that would be useful in the development of a compound. Provided in this invention are additional genetic pathways implicated in the disease process or response to candidate therapies. Variation in these genes may account for the observed variability in treatment response. Exemplary variations in the candidate genes are provided in Tables 12-17 and 18-23. The Detailed Description above describes how one skilled in the art would identify a candidate gene or genes, identify sequence variances, stratify patients, design clinical trials, and obtain regulatory approval of a pharmacogenetic test for optimal responders to an ALS treatment.Gene pathways including most preferably, but not limited to, those genes that are listed in the gene pathway Table 2, and pathway matrix Table 7 and discussed in Section V. below are candidates for the genetic analysis and product development strategies described above.

Advantages of Pharmacogenetic Clinical Development of Agents for ALS

In view of the limitations of present therapy, the advantages of an ALS clinical development program that includes genetic stratification of patients in the analysis of response to candidate therapeutic interventions are numerous. First, it may be possible to identify a subpopulation that responds to a treatment at a higher rate than the whole ALS population. This would address the demonstrated disinclination of ALS patients to expose themselves to therapies of limited effectiveness. It might also allow regulatory approval of therapies that do not produce a sufficient response in the unstratified population to justify approval. Second, it may be possible to identify patients who respond to a treatment only at higher doses than most patients require, or respond preferentially to an altered dosing route or schedule. Such customization of therapy to individual genetic and biochemical differences may allow a higher overall response rate to be achieved, without requiring totally empirical dose adjustment in each patient. Third, it may be possible to identify patients in whom side effects are likely to occur. Such patients could be offered alternative treatments. It is also worth noting that the type of benefit afforded by drugs such as RILUTEK®- a slowing of deterioration – will likely be most useful if the drug is started very early, before large numbers of neurons are gone. However the long term prophylactic use of medicines in well, or nearly well, individuals entails a different cost-benefit analysis than in already sick individuals. Identification of patients that respond well to early neuroprotective therapy may be aided by the analysis of genetic determinants of treatment response.

Additional uses of genetic stratification in clinical development have been described above.

As an example of a candidate gene with DNA sequence variances potentially relevant to drug efficacy, safety, or both consider the glutamate aspartate receptor NMDA 2C, a member of the glutamate pathway. Described in this application are novel NMDA 2C DNA sequence variances that the inventors have recognized may affect response to drugs. (Diseases in which the glutamate pathway is likely to play a role are summarized in Table 7) Six DNA sequence variances have been identified in the NMDA 2C gene, five of which alter the encoded amino acid sequence. Several of the amino acid variances are nonconservative, including phenylalanine-valine, glycine-arginine and arginine-serine (see Table 13 for details). Seven DNA sequence variances are described in the NMDA 2A receptor (Table 13). The effect of one or more of these genetic polymorphisms on efficacy or safety of an ALS treatment could be tested in a clinical trial. For example, the goal could be optimization of patient selection for glutamate channel antagonist therapy of ALS by determining whether an ALS patient has a NMDA 2A or 2C receptor genotype against which a glutamate antagonist is more effective or safer.

Similarly, for genes belonging to the other pathways relevant to treatment of ALS (see tables 2 and 7) and polymorphisms in those genes (tables 13 and 19) a strong argument can be made that said polymorphisms (or sets of polymorphisms, or haplotypes) may affect efficacy or safety of drugs active against ALS, including, but not limited to, drugs listed below in Table 25 and related compounds. The candidate genes include, but are not limited to, modulators of glutaminergic, serotonergic. GABAergic, melatonergic and opiate pathways, as well as calcium channels, cytokines, factors that mediate growth, differentiation and apoptosis, the coagulation cascade, second messenger systems, detoxification genes, particularly relating to superoxide, protein degradation and cytoskeleton genes.

*V. Therapeutic Strategies for ALS*

The etiology of most ALS cases is unknown but may involve autoimmune responses, for example to calcium channels, injury due to excessive excitotoxic stimulation (especially via aspartate, glutamate and GABA receptors), impaired clearance of free radicals, imbalance of neurofilament turnover or possiblly viral mediated destruction of motor neurons (e.g. herpes virus). A number of drug development programs are aimed at these postulated pathophysiologic mechanisms. For example, there are candidate therapeutic agents that down modulate immune reactivity, block or dampen excitatory neurotransmitter signalling, alleviate free radical injury, and interfere with a hypothesized viral infection of motor neurons.

Beyond the specific mechanisms of action enumerated above, there are many compounds in development that are intended to halt, retard, or prevent neural cell degeneration, or promote neural cell regeneration. Many such compounds are in clinical development programs for multiple neurological diseases. For example, gabapentin is a compound with complex and incompletely understood pharmacology, but its shows anticonvulsant, antinociceptive, anxiolytic and neuroprotective activity in animal models. In ALS animal models gabapentin prevents neuronal death. One of its actions may be inhibition of glutamate synthesis by branched-chain amino acid aminotransferase (BCAA-t). Other compounds in development for ALS target proteins involved in growth control and differentiation, protein processing, intracellular second messenger cascades and cytoskeletal proteins (see 25 below for specific compounds and Table 1 for the candidate genes that may affect response to those compounds).

Below in Table 25 the therapies in development for ALS catagorized by mechanism of action. The listed candidate therapeutic intervention response in patients with ALS may be affected by polymorphisms in genes as described above in the Detailed Description.

**Example 11**

Dementia

*I. Description of Dementia*

Dementia is a general term for mental deterioration. clinical state characterized by a significant loss of function in multiple cognitive domains, not due to an impaired level of confusion. Diagnosis of dementia requires 1) assessment of an individual's current level of cognitive function with the ability to compare to past intellectual function, and 2) documenting a decline in intellectual function by serial examinations over time. A comprehensive, reliable, and universally accepted clinical classification of the clinical and neuropathological characteristics of senile dementia has been described. However, definitive diagnosis is obtainable only with pathological findings upon autopsy. Based upon these diagnoses, there are an estimated 4 million Americans with Alzheimer's disease (AD) and 10 million Americans with dementia of all types.

Besides AD, there are categories of dementia that include vascular dementia, lewy body disease, frontal lobe dementia, mixed dementia, and post-traumatic dementia. A number of different diseases or conditions are characterized by or involve loss of cholinergic function and/or defects in neuronal remodeling repair and may result in clinical symptoms of dementia. Among these are diseases such as Alzheimer's disease (AD), Huntington's disease, Parkinson's disease, and

amyotrophic lateral sclerosis (ALS). Dementia can further be a complication of the following: depression, drug intoxication, metabolic disorders, normal pressure hydrocephalus, subdural hematomas, and cerebrovascular insufficiency.

II. *Current Therapies for Dementia*

Current therapies for the treatment of dementia include enhancement of cortical cholinergic function. In general, approaches to replacement of cholinergic function can be characterized as either: 1) therapies that compensate for existing damage; and 2) therapies that halt, retard, or prevent cerebral damage. Ideally, a therapy targeting both mechanisms could potentially reverse existing damage. There are two broad mechanisms to enhance cerebral cholinergic function; 1) to block metabolism of acetylcholine via an acetylcholinesterase inhibitor, or 2) agonists at muscarinic or nicotinic receptors.

Acetylcholinesterase inhibitors have recently been approved for the use in patients with mild to moderate Alzheimer's disease. These agents (donepezil, Tacrine) selectively inhibit the acetylcholinesterase enzyme and increases levels of cortical acetylcholine. In randomized controlled clinical trials, donepezil was shown to improve both cognitive performance and global functioning. The improvements are modest and may not be apparent until up to three months after commencement of treatment.

*III. Limitations of Current Therapies for Dementia*

Despite the introduction of pharmacologic agents for the treatment of dementia, the mainstay of therapeutic management continues to be education, and support for caregivers, and treatment of complications. This is in part because the available acetylcholinesterase inhibitor (donepezil) has limited efficacy and has undesirable side effects. Thus, the clinician is faced with the dilemma of limited therapeutic alternatives and weighing the benefits against the side effects.

Limitations of Acetylcholinesterase Therapy due to Low Efficacy

Acetylcholinesterase inhibitors have limited efficacy; only a fraction (modest improvement in 40-50%) of patients respond to therapy. The extent and progression of loss of cortical cholinergic neurons limit the therapeutic benefit of acetylcholinesterase inhibitors. Long-term benefit of inhibition of acetylcholinesterase activity is unproven. Further, there is no clinical evidence supporting the use of acetylcholinesterase inhibitors in the prevention of AD or in the treatment of more severe stages.

An additional efficacy concern of the acetylcholinesterase inhibitor is the latent period before demonstrable clinical benefit. In the same period there may be concurrent neurodegeneration. Thus, the clinician has limited therapeutic alternatives, the patient may have limited response to therapy, and the disease progresses. In many cases, medical management of dementia is reduced to treatment of complications or supportive care.

Limitation of Acetylcholinesterase Therapy due to Toxicity or Undesired Side Effects

Toxicities associated with the use of acetylcholinesterase inhibitors are 1) vagotonic effect on the myocardium resulting in bradycardia and complicationsof other myocardial syndormes, 2) gastrointestinal complications such as nausea, vomiting, diarrhea, 3) lowering of seizure threshold (since seizures can be a complication of AD, this side effect may be confused with the progression of the disease).

Other acetylcholinesterase inhibitors have been shown to have a severe hepatotoxic effect, those products have been removed from the market or clinical development programs.

*IV. Impact of Genotyping on Drug Development for Dementia*

As previously indicated, the pathways and genes emphasize the relationship with Alzheimer's disease. In connection with the development of Alzheimers, it had been found that the presence of the ApoE4 allele was associated with an earlier development of the disease than other alleles, and further was associated with a decreased response to present acetylcholinesterase inhibitors, such as tacrine. The ε4 allele of Apolipoprotein E (ApoE) is a well-established risk factor for late onset Alzheimer's disease. The work of Poirier (1995) and Farlow (1998) suggests there are significant interactions between sex, ApoE genotype, and therapeutic response (ADAS-Cog scores) to the acetylcholinesterase inhibitor tacrine, with the ε4 allele generally associated with poor response and the effect being more notable in women than in men. ApoE is only part of the brain lipid transport pathway, however, and the interaction of allelic variation at other components of this pathway with drug response can also contribute to variation in therapeutic responses.

Sequence variance in the butyrylcholinesterase (BCHE) gene has been found to correlate with the development of Alzheimer's disease, as well as with treatment efficacy of both cholinomimetic and non-cholinomimetic drug therapies. In this case, the presence of at least one BCHE-k allele is predictive of the development of

Alzheimer's disease and is negatively correlated with treatment efficacy of tacrine (a colinestrase inhibitor) and an experimental vasopressinergic drug (a non-cholinomimetic drug). The BCHE-k allele has a point mutation at nucleotide 1828 (a G to A subsitution) which results in an ala539thr change. This polymorphism can be readily detected by PCR amplifying a region surrounding the variance site and sequencing the amplification product to determine the nucleotide at the particular site.

A group of patients was treated with an experimental vasopressinergic drug (n = 91) and compared to patients administered a placebo (n = 108) without segregation or stratification by BCHE or other allelic status. As evaluated using the Mini Mental State Examination (MMSE) over a twelve-week treatment period, no statistically significant improvement was shown for the treatment group. However, when the treatment group was stratified according to the presence or absence of a BCHE-k allele, those patients without such an allele showed a statistically significant improvement while those having at least one of the BCHE-k alleles did not. Thus, the analysis provides an example of a gene where a patient sub-population was identified where a treatment showed a positive response even though no such positive response was found for the overall patient population. Indeed, those patients not having a k-allele are approximately three times more likely to respond to the vasopressinergic drug than are patients having at least one k-allele.

The response of Alzheimer's disease patients treated with the cholinomimetic drug, tactrine, was also determined. Similar to the above, the MMSE test was utilized as an indicator of a positive response. The positive response rate was approximately two-fold higher in those patients not having a k-allele than in those patients having at least one k-allele.

In addition, it was found that the presence of either or both of a BCHE-k allele and an apoE-4 allele was positively correlated with the development of Alzheimer's disease. For example, in patients over 75 years of age, the odds ratio of a patient having a BCHE-k allele was 2.3, the odds ratio for having a apoE-4 allele was 2.0, and the odds ratio for the joint occurrence of both alleles was 17.5. Thus, the BCHE-k allele is an example where the presence of a variant allele is negatively correlated with the efficacy of treatment with drugs from multiple drug categories, and which is further postively correlated with the development of a particular disease. Thus, the variance status of such a gene is useful both as a prognostic tool for disease risk, as well as for identifying likely drug responders versus non-responders for drug development and/or treatment selection.

The evidence that a variance in a gene involved in a pathway that affects drug response in patients with dementia, indicates and supports the theory that there

is a likelihood that other genes have similar qualities to various degrees. As drug research and development proceeds to identify more lead candidate therapeutic interventions for dementia, there is possible utility in stratifying patients based upon their genotype for these yet to be correlated variances. Further, as described in the Detailed Description, methods for the identification of candidate genes and gene pathways, stratification, clinical trial design, and implementation of genotyping for appropriate medical management of a given disease is easily translated for dementia. As described below in section V. below there are likely gene pathways as are those that are outlined in the gene pathway Table 2 and matrix Table 7.

Advantages of Pharmacogenomic Clinical Development of Therapies for Dementia

The advantages of a clinical research and drug development program that includes the use of polymorphic genotyping for the stratification of patients for the appropriate selection of candidate therapeutic intervention includes 1) identification of patients that may respond earlier to therapy, 2) identification of the primary gene and relevant polymorphic variance that directly affects efficacy, safety, or both, 3) identification of pathophysiologic relevant variance or variances and potential therapies affecting those allelic genotypes or haplotypes, and 4) identification of allelic variances or haplotypes in genes that indirectly affects efficacy, safety or both.

Based upon these advantages, designing and performing a clinical trial, either prospective or retrospective, which includes a genotype stratification arm will incorporate analysis of clinical outcomes and genetic variation associated with those outcomes, and hypothesis testing of the statistically relevant correlation of the genotypic stratification and therapeutic benefits. If statistical relevance is detectable, these studies will be incorporated into regulatory filings. Ultimately, these clinical trial data will be considered during the approval for marketing process, as well as, incorporated into accepted medical management of dementia.

By identifying subsets of patients with mild to moderate dementia that respond earlier to drugs or agents or experience enhanced efficacy, optimal candidate therapeutic interventions may reduce the period of time prior to relief of cognitive impairments. Appropriate genotyping and correlation to dosing regimen would be beneficial to the patient, caregivers, medical personnel, and the patient's loved ones.

Optimization of cholinomimetic mediated therapy of dementia further demonstrates the utility of selection of a potential dementia patient that has a predisposing genotype in which selective cholinomimetic are more effective and or are more safe. In considering an optimization protocol, one could potentially

predetermine variance or variances within the muscarinic cholinergic receptor, nicotinic cholinergic receptor, modulatory mechanisms of cholinergic neurotransmission, or cholinergic receptor mediated intracellular mechanism of action that is preeminently responsible for drug response. By embarking on the previously described gene pathway approach, it is technically feasible to determine the relevant genes within such a targeted drug development program for dementia.

*V. Description of Mechanism of Action Hypotheses for Future Drug Development for the Therapy of Dementia*

Drug development programs for the identification of novel drug or candidate therapeutic interventions are aimed at the underlying pathophysiologic mechanisms of the disease leading to clinical signs and symptoms of dementia. Current hypotheses include, but are not limited to, therapeutic development in one of the following areas: 1) replacement of cholinergic function, 2) acetylcholine pathway: biosynthesis, secretion, degradation, reuptake, and receptor binding, 3) CNS lipid transport/membrane repair pathway and gene identification, 4) inflammatory mediators, e.g. prostaglandin, prostacylin, and thromboxane pathway, and 5) constituents of AD lesions and AD genes. These are described in detail below.

A. Therapeutic Approaches for Replacement of Cholinergic Function

Because dementia is apparently related to a loss of cholinergic function in the neocortex and forebrain arising from death or atrophy of basal forebrain cholinergic neurons, replacement of cholinergic function has been shown to have therapeutic benefit. In general, novel approaches for the replacement of cholinergic function can be characterized as either: 1) therapies that compensate for existing damage; and 2) therapies that halt, retard, or prevent cerebral damage.

B. Therapeutic approaches that compensate for existing damage

The therapeutic approaches that may compensate for existing damage include modulating cholinergic deficit, modulating other neurotransmitter deficits, modulating immune or inflammatory mechanisms of neural damage, and modulation of metabolism of specific neurotransmitters. Although these novel therapies are aimed at existing or damage yet to occur, the underlying course of the disease will remain.

Potential therapies for the compensation for cholinergic deficit are 1) increase presynaptic production of acetylcholine, 2) enhance release of acetylcholine, 3) stimulate choline reuptake, 4) selective muscarinic agonists, 5)

anticholinesterase inhibitors, 6) mixed action anticholinesterases and muscarinic receptor ligands, and 6) nicotinic receptor agonists.

Potential therapies for the compensation for modulating other neurotransmitters are 1) selective NMDA agonists, and 2) other disorders of neurotransmitter function.

Potential therapies for the compensation for modulating immune or inflammatory mechanisms of neural damage are 1) antiinflammatory agents that suppress inflammation and 2) inhibition of amyloid precursor protein (APP) degradation.

Potential therapies that compensate for monoamine metabolite deficits are agents that affect monoamine oxidase type B enzyme activity, therapy for behavioral symptoms of neurotransmitter function in dementia, and compensate for immune or inflammatory mechanisms involved in neural cell destruction.

C. Therapeutic approaches that halt, retard, or prevent cerebral damage

In general therapeutic approaches that halt, retard, or prevent cerebral neural damage are currently either growth factors or modulation of the deposition of aberrent pathological depositions of metabolic by-products. These approaches include promotion of the growth and regeneration of cholinergic neurons and generally include growth factors that act on neurons, neural precursors, or glial cells. Growth factors includebut are not limited to nerve growth factor (NGF), brain-derived growth factor (BDGF), neurotrophins, and leukemia inhibitory factor (LIF).

Prevention of amyloid plaque deposition includes modulation of APP gene expression, prevention of the development of amyloidogenic peptide, inhibition of amyloid aggregation/secretion, and APP antagonists. Prevention of the formation of neurofibrillary tangles includes modulation of the phosphorylation of tau proteins.

D. CNS Lipid Transport/Membrane Repair Pathway and Gene Identification

Brain Apolipoproteins: The six apolipoproteins known to be expressed in the brain are listed below. They are present on the surface of three major types of lipoproteins, one class enriched in A-I, but also containing most of the D,E, and J protein in the brain; one class composed principally of E with minor amounts of A-I, A-IV, D, and J; and a third minor class containing the majority of A-IV. Variation in the structure or expression of these apolipoproteins can modulate lipid transport and brain remodeling.

Lipoprotein Receptors: Six brain receptors for lipoproteins have been identified in man. These include the low density lipoprotein receptor (LDL-R), the LDL receptor-related protein (LRP), the very low density lipoprotein receptor

(VLDL-R), and the class A macrophage scavenger receptor, all of which are also expressed outside the brain. Two new protein with LDL receptor-like domains have recently been identified in human brain: Apolipoprotein E receptor type 2, and the SorLA-1 receptor. Alterations in the structure or expression of those receptors can affect binding of ApoE alleles (ApoE2, for example, has reduced affinity for the LDL receptor), and more generally will modulate the biology of lipid transport.

Lipoprotein docking and lipid mobilization: Heparin sulfate proteoglyans (HSPG) are responsible for initial binding of ApoE-bearing lipoproteins to cells. Removal of HSPGs with heparinase blocks binding, even in the presence of receptor (LDL-R or LRP). Therefore variations in biosynthetic enzymes of the HSPG pathway will influence lipoprotein uptake. Lipid hydrolysis by cholesterol ester transfer protein (CETP) effects the transport of lipids from lipoproteins into cells.

Cholesterol Metabolism: Acyl CoA:cholesterol acyltransferase and HMG CoA reductase are responsible for the metabolism of cholesterol, therefore variations in the metabolic pathway of cholesterol will influence availability of cholesterol.

Hormonal control of lipoproteins and lipoprotein receptors: The expression of lipoproteins and their receptors is under hormonal control. Clinical studies of tacrine for Alzheimer's disease have also shown reduced incidence of AD in women taking estrogen supplements post-menopausally. Therefore, variation in hormone levels, hormone receptors, or hormone receptor signaling pathways will modulate response to acetylcholinesterase inhibitors, e.g., by affecting lipid transport and cholinergic remodeling or by other means. Hormone receptors that bind their physiologic ligand within the cytoplasm then become activated and cross the nuclear membrane include but are not limited to growth hormone, prolactin, estrogen, retinoic acid receptor, thyrotropin releasing hormone. Associated transcriptional co-activators include but are not limited to SRC-1, SRC-2 (TIF-2), SRC-3 (p/CIP:AIB1), P/CAF, CBP, E6-AP, TRIP230, SMRT, SRA, and N-CoR.

E. Prostaglandin, Prostacylin, and Thromboxane Pathway

Inflammatory mediators, and in particular the products of arachidonic acid metabolism, play a role in the development of AD neuropathology.

There are several lines of evidence supporting the role of inflammatory or immunological processes in the pathogenesis of Alzheimer's disease. First, neurodegeneration in AD is accompanied by manifestations of immune reaction including activation of the complement cascade, accumulation and activation of microglia and presence of inflammatory cytokines and acute phase reactants in tissue of AD brains. Second, epidemiological studies suggest that use of non-steroidal anti-inflammatory drugs (NSAIDs) delays the clinical expression of

Alzheimer's disease. The development of selective COX inhibitors has led to renewed interest in the therapeutic potential of NSAIDs in AD.

Arachidonic acid formation pathway genes include phospholipase A2, phospholipase C β3, and diacylglycerol lipase. PGG2 formation pathway genes include cyclooxygenase I, cyclooxygenase II. PGH2 formation pathway genes include PGG2 reductase. PGH2 metabolizing enzymes include PGH2 reductase, PGD2 reductase, PGH-PGE isomerase, and thromboxane A2 synthase. Receptors include PGF1a receptor, PGD2 receptor, PGE2 receptor, PG12 receptor, and thromboxane A receptor. Exemplary variances for genes above are shown in Tables 13 and 19.

F. Constituents of Alzheimer's Disease Lesions and AD Genes

The relative contribution of different pathogenetic mechanisms to the development of AD in specific patients can affect the degree of cholinergic impairment and hence the response to acetylcholinesterase inhibitors.

There is clear evidence that different pathogenetic mechanisms affect the onset and rate of progression of AD. The possible effects of such are several lines of evidence supporting the role of inflammatory or immunological processes in the pathogenesis of Alzheimer's disease. First, neurodegeneration in AD is accompanied by manifestations of immune reaction including activation of the complement cascade, accumulation and activation of microglia and presence of inflammatory cytokines and acute phase reactants in tissue of AD brains. Second, epidemiological studies suggest that use of non-steroidal anti-inflammatory drugs (NSAIDs) delays the clinical expression of Alzheimer's disease. The development of selective COX inhibitors has led to renewed interest in the therpaeutic potential of NSAIDS in AD. Pathway genes include Tau protein, amyloid precursor protein, presenilin 1, and presenilin 2.

In Tables 13 and 19, there are listings of candidate genes and specific single nucleotide polymorphisms that may be critical for the identification and stratification of a patient population diagnosed with dementia based upon genotype. Current pathways that may have involvement in the therapeutic benefit of dementia include, but are not limited to, glutaminergic, serontonergic, dopaminergic, adrenergic, cholinergic, histaminergic, purinergic, GABAergic, glycinergic, nitric oxide, peptide protein processing, opiates, cholecystokinin, corticotropin releasing factor, thyroid stimulating hormone, somatostatin, adrenocorticotropic hormone, vasoactive intestinal peptide, calsium or potassium channels, prostaglandin, cytokines, estrogen, clot formation, hemostasis, oxygenstress, mitochondrial

maintenance, protein maturation and degradation, second messenger cascade, growth, differentiation and apoptosis, cytoskeleton, secretion, amyloid processing, and lipid transport or metabolism gene pathways that are listed in Tables 1-6, 12-17 and 18-23. One skilled in the art would be able to identify these pathway specific gene or genes that may be involved in the manifestation of dementia, are likely candidate targets for novel therapeutic approaches, or are involved in mediating patient population differences in drug response to therapies for dementia.

A sample of therapies approved or in development for preventing or treating the progression of symptoms of dementia currently known in the art is shown in Table 27. In this table, the candidate therapeutics were sorted and listed by mechanism of action. Further, the product name, the pharmacologic mechanism of action, chemical name (if specified), and the indication is listed as well.

Based upon these varying approaches there are many products in development for dementia. In Table 27 below lists current therapies that are in development for U.S. marketing approval. Identification of genes in specific pathways and the link to specific agents or drugs may be useful to conduct clinical trials and achieve higher degrees of safety and efficacy. The listed candidate therapeutic interventions response in patients with dementia may be affected by polymorphisms in genes as described above.

**Example 12**

Depression

*I. Description of Depression*

Major depression is a psychiatric disorder distinguishable from normal grief, sadness, and disappointment as well as the dysphoria and demoralization often associated with medical illness. Depressive disorders are characterized by abnormally long term depressed mood and may be accompanied by delusions and hallucinations. Individuals suffering from depression have feelings of despair and intense sadness, exhibit mental slowing and loss of concentration, are preoccupied with pessimistic worry and inner self, and are agitated and tend toward self-deprecation. In some depressive disorders, mania is present usually in episodic intervals and in these cases depressed mood is replaced with feelings of grandiosity and may be accompanied by incoherent speech. Clinically, unipolar or bipolar depression are terms used to describe the two broad categories of depressive disorders characterized by the absence or presence of episodic mania, respectively.

*II. Current Medical Management Strategies for Depression*

Unipolar Depression

Depression is a wide-spread disease that requires improved therapeutic alternatives to the conventional agents that have been available since the 1960s. Current therapeutic candidates of unipolar or bipolar depression are as follows: tricyclic antidepressants, tetracyclic antidepressants, lithium, monoamine oxidase (MAO) inhibitors, electroconvulsive therapy (ECT) and atypical agents such as PROZAC®, WELLBUTRIN®, and trazodone.

Bipolar Depression

Despite the difficulties of medical management of bipolar depression, advances have changed therapeutic outcomes. Therapies such as lithium, valproate, and carbamazepine, clozapine, and ECT have made a positive impact on the patient outcomes. Further, the importance of psychosocial issues for understanding patients illnesses and factors affecting treatment compliance are more fully realized.

For bipolar depression, mood stabilizers are the first line therapy and include: lithium, valproate, and carbamazepine. Adjunct therapies are used for the treatment of agitation, insomnia, or aggressive behaviors and include benzidiazepines and antipyschotics. ECT is useful as an alternative therapy in patients who are pregnant or are trying to conceive, unresponsive to standard therapy, unable to tolerate first line therapies, or are refractory to first line or adjunct therapies. ECT has been shown to be effective as stated above, as well as 54% effective in refractory patients.

There are additional therapies that have been used for the treatment of bipolar depression. For example, off-label use of clozapine, Ca++ channel antagonists, gabapentin, and lamotrigine in diagnosed bipolar patients have been demonstrated to be effective at stabilizing mood. Gabapentin, has a higher safety profile during pregnancy, but has side effects of ataxia, fatigue and somnolence. Lamotrigine, by effectively lowering glutamine release is effective at stabilizing mood, but is associated with dizziness, headache, double vision, somnolence, headache, and rash. Other medications include valproate for euphoric mania, valproate for dysphoric mania or mixed mania, and clozapine with lithium or valproate for patients with rapid-cycling episodes.

*III. Limitation of Current Therapies for Depression*

Frequently, depression is undiagnosed and if detected, treatment often is inadequate. Therapy of depression is associated with undesirable side effects and/or simply fails to adequately manage the symptoms of the condition. Thus, there is a

need for ongoing improved development of antidepressant therapeutic alternatives to the currently available products.

Limitations of Current Therapies for Unipolar Depression

Although these agents or therapies are efficacious (e.g. 80% improvement following ECT; lithium effectively prevents relapses in 60% of patients) there are significant limitations to their use and are 1) the onset of action of antidepressant drugs is latent, 2) responsivity and efficacy is not uniform, 3) long-term treatment can lead to symptoms of drug resistance, 4) there is perceived inhibition of creativity and decreased energy, and 5) there are patients with refractory depression with no therapeutic alternatives.

Limitations of Current Therapies for Bipolar Depression

Bipolar depression patients have additional therapeutic concerns as compared to unipolar depression patients. For bipolar patients there is the added difficulty of treating depression episodes. The efficacy of antidepressants is not well founded or documented in bipolar depression. Further, antidepressants have been documented to induce manic or hypomanic symptoms. Therefore, mood stabilizers are the first line therapy with adjunct therapies during manic or depression episodes.

An additional therapeutic issue associated with bipolar patients is that many comorbid psychiatric disorders occur within the same patient not only hindering a diagnosis, but also therapy. For example, substance abuse disorders, panic disorders, obsessive-compulsive disorders, and impulsive control disorders are often present and potentially mask symptoms of bipolar depression.

*IV. Impact of Pharmacogenomics on Drug Development for Depression*

There are two genes that have been described having polymorphisms that affect antidepression drug response, the serotonin transporter gene and the angiotension converting enzyme that affects the metabolism of substance P. These two examples are described below.

The Serotonin Transporter Gene

The serotonin transporter gene (5-HTT) polymorphism provides an example of a recessive SNP polymorphism in the non-coding region with an impact on inefficacy of a 5-HTT selective drug.

The serotonin transporter (5-HTT) plays a critical role in the termination of the serotonin (5-HT) neurotransmission and represent the prime target for selective

serotonin reuptake inhibitors (SSRIs). A functional polymorphism in the transcriptional control region upstream of the 5-HTT coding sequence has been reported. It consists of a 44 –base pair insertion (long variant) or deletion (short variant). It has been demonstrated that the long (l) and short (s) variants of this 5-HTT gene-linked polymorphic region had different transcriptional efficiencies. *In vitro* studies showed that the difference in 5-HTT mRNA synthesis result in different 5-HTT expression and 5-HT cellular uptake (Lesch et al. Science 1996 274:1527-153). Recently, it has been shown that an SSRI (fluvoxamine) efficacy in delusional depression seems to be related to allelic variation within the promoter of the 5-HTT gene (Smeraldi et al. Mol. Psychiatry 1998; 3:508-511). Both homozygotes for the long variant (l/l) of the 5-HTT promoter and heterozygotes (l/s) showed a better response to fluvoxamine than homozygotes for the short variant (s/s). Interestingly, the addition of pindolol (a mixed adrenoreceptor and 5-HT1A antagonist) has been proposed as an augmentation therapy for non-responders or partial responders to SSRIs, and it appears that in the group treated with fluvoxamine plus pindolol all the genotypes acted like l/l treated with fluvoxamine alone. This supports the hypothesis that the effect of pindolol is related to its ability to block 5-HT1A autoreceptors, thus preventing a negative feed-back of 5-HT at somatodendritic level. Furthermore, the activation of 5-HT1A autoreceptors could modulate the clinical effect of the SSRIs-induced 5-HTT blockade.

The 5-HTT polymorphism represents an example of a gene allelic variance that affects the transcriptional control, and ultimately, the amount of available transporter protein. In these cases, the gene product concentration or protein availability affects the function of the native mechanism and ultimately the ability of the drug to intervene with physiological function. One skilled in the art, upon utilizing the techniques described in the detailed description, would be able to identify known variances within a candidate gene, provide a diagnostic test to identify individuals with that variance or variances, group the individuals based upon the identified genotype, and design and implement a clinical study to test the effect a candidate drug has on the the groups. In this example, the allelic differences may affect transcriptional or translational control of the 5-HTT gene. A skilled practitioner will be able to utilize the techniques known in the art to determine the effects of a variance or variances within a gene promoter region to be able to study the impact those allelic differences have on the safety or efficacy of SSRIs or any other candidate drugs affecting the 5-HT pathway. Further, this example underscores the ability of a skilled practitioner to be able to utilize methods known in the art to design a pharmacogenomics clinical trial when the allelic difference is within the gene promoter region.

The Angiotensin Converting Enzyme Gene and Substance P

The localization of substance P in brain regions that coordinate stress responses and receive convergent monoaminergic innervation suggested that substance P antagonists might have psychotherapeutic properties. Similar to clinically used antidepressant and anxiolytic drugs, substance P antagonists suppress isolation-induced vocalizations in guinea pigs. In a placebo-controlled trial in patients with moderate to severe major depression, robust antidepressant effects of the substance P antagonist MK-869 were consistently observed. In preclinical studies, substance P antagonists did not interact with monoamine systems in the manner seen with established antidepressant drugs. These findings suggest that substance P may play an important role in psychiatric disorders.

Substance P is highly metabolized by ACE (angiotensin converting enzyme) which is a good actual example of pharmacogenetics: It has a high allele frequency in normal individuals (D: 34%, I: 66%) and there are clinical studies clearly demonstrating its impact on ACE inhibitors.

Moreover, it has been shown that DD homozygous patients (11%) have a higher brain level of substance P than II homozygous patients (43%), with an intermediate level for heterozygous patients (46%).

Using results of the initial phase II trial, we expect that a substance P antagonist will have more impact on patients with high brain level of substance P (actually, the DD patients who are more at risk for affective disorders). As measure of response rate, starting with the standard measure of response defined as $\geq$50% change from baseline to week 6 in total HAM-D21 score, 54% of the patients improved with MK-869 and 28% patients improved with placebo in the phase II trial.

In a recent clinical trial of MK-869 versus placebo, a similar response rate was observed for both groups (54% and 48% respectively). If the ACE variance is considered as a dominant SNP with regard to substance P metabolism, calculation of an unequivocal positive response rate in the DD subgroup (i.e., 100%) would require an equally similar response rate in the II subgroup, while assuming the DI subgroup response rate remains similar to placebo (i.e., 48%). In this case, MK-869 would be positive (100%) only in a fraction of the patients, e.g., one out of every five.

Approximately 25% of the responders should be DD homozygous; if not, the hypothesis is not valid. Then, if 25% are DD, the number of patients included in the failed trial should be enough to see a statistically significant difference between the DD subgroup and other patients, since we would need at least 56 patients to test for such a high relative risk (100% / 48% = 2).

This approach exemplifies the utility of high allele frequency polymorphisms. Further, when the treatment is not efficacious for all individuals (i.e. response rates vary between treatment groups is less than 15%) the allele frequency of a potentially interacting recessive SNP polymorphism should be relatively high (e.g. from 30% for a 15% difference in response rate to 60%). This corresponds to 16% or less of total patients (see example 18 and table below).

The evidence that a variance in a gene involved in a pathway affects antidepressive drug response, indicates and supports the idea that other genes have similar qualities to various degrees. As drug research and development proceeds to identify more lead candidate therapeutic interventions for depression, there is utility in stratifying patients based upon their genotype for these yet to be correlated variances. Further, as described in the Detailed Description, methods for the identification of candidate genes and gene pathways, stratification, clinical trial design, and implementation of genotyping for appropriate medical management of a given disease is easily utilized for depression. As described below in section V. below there are likely gene pathways such as those outlined in the gene pathway Table 2 and matrix Table 7.

Optimization of adrenergic control or ion channel modulation mediated therapy of epilepsy further demonstrates the utility of selection of a potential epilepsy patient that has a predisposing genotype in which selective adrenergic or agents are more effective and or are safer. In considering an optimization protocol, one can potentially predetermine variance or variances within the adrenergic receptor, ion channel or ion channel mediated mechanisms of neurotransmission, or adrenergic receptor mediated intracellular mechanism of action that is preeminently responsible for drug response. By embarking on the previously described gene pathway approach, it is technically feasible to determine the relevant genes within such a targeted drug development program for depression.

A sample of therapies in development for preventing or treating the progression of symptoms of depression currently known in the art is shown in table 28. In this table, the candidate therapeutics were sorted and listed by mechanism of action. Further, the product name, the pharmacologic mechanism of action, chemical name (if specified), and the indication is listed as well.

## *V. Mechanism of Action Hypotheses for Novel Therapies for Depression: Utility of Genotyping*

Unipolar Depression

Unfortunately, to date the biological mechanism of major unipolar depression is unclear. However, studies of endocrine systems, neurotransmission, and neuroelectrophysiology have provided the basis for the generation of pathophysiologic hypotheses. These hypotheses have been supported by clinical data stemming from the success of conventional treatment of depression.

One such hypothesis is that there is pituitary-hypothalamic dysfunction in depressed patients. It has been observed that depressed patients commonly have elevated levels of cortical steroids in their urine and blood. Further, 50% of the patients with clinical depression will not secrete cortisol when subjected to the dexamethasone suppression test. Additionally, thyrotropin releasing hormone (TRH) stimulation of thyrotropin stimulating hormone (TSH) release is aberrant in depressed patients without an alteration of serum T3 or T4 concentrations and growth hormone, prolactin, gonadal hormones, corticotropin releasing factor (CRF), and melatonin have diminished physiologic responses.

Another hypothesis of the biological dysfunction of depression is that there is a neurotransmitter dysfunction due to a catecholamine-indolamine imbalance. This theory postulates that there is a required level of catecholamines and receptor sensitivity required for normal mood. In depression, there may be aberrant receptor insensitivity, depletion of amines, or a depletion of their synthesis or storage that leads to depression. Supporting this theory is that monoamine oxidase inhibitors increase the availability of catecholamines and indolamines and have been used clinically for the management of depression.

The cholinergic neurotransmitter system has been implicated in the manifestation of depression. In has been postulated that there is an imbalance of adrenergic and cholinergic control of neural transmission in patients with depression.

Electrophysiologic studies have shown that patients with depression have altered rapid eye movement (REM) sleep patterns, i.e. shortened REM latency, than non-depressed patients. Other studies have documented a correlation of the circadian rhythm and precipitation of depressive episodes during autumn and winter months and diminished ambient light during those times during the year.

In each of the theories posited and described above, satisfactory conclusions are limited. Conventional therapy of depression with tricyclic antidepressants has demonstrated that this treatment affects more than one neurotransmitter system due to either modification or alteration of the regulation of neurotransmitter receptors signaling pathways rather than acting solely at neurotransmitter receptor binding.

Novel therapies of unipolar depression include venlafaxine and mirtazapine. Both of these compounds show promise in clinical trials for the treatment of

depression. Venlafaxine is a mixed serotonergic and noradrenergic reuptake inhibitor. Mirtazapine has noradrenergic and serotonergic antidepressant mechanism of action. These two products have what looks to be superior action over tricyclic antidepressants or selective serotonergic inhibitors (SSRIs).

Bipolar Depression

Theories for the mechanism have been described. In one model, electrophysiological kindling and behavior sensitization underlie bipolar disorders and further increasing frequencies of episodes over time. In another model, there appears to be a desynchronization of circadian rhythm in bipolar patients.

As for depression, the catecholamine hypothesis presumes that mania is due to an excess of catecholamines and depression is due to their depletion. Noradrenergic and dopaminergic dysfunction have both been linked to depression. In both cases of dysfunction, there appears to be causal links, i.e. aberrant noradrenergic neurotransmission and L-dopa induced hypomania among bipolar patients, respectively. Amphetamines can produce hypomania in bipolar patients and dopaminergic antagonists are effective for severe mania.

The serotonergic hypothesis generalizes that low serotonergic transmission is responsible for mania and depression because low serotonergic inputs may result in defective neuromodulation. Other hypotheses include neurotransmitters, enzymes, neuropeptides, and theories involving endocrine and immunological systems. As in many other complex disorders of psychological function, these models fall short of adequately describing the disturbance. Future studies and drug development may provide insights to refined biological mechanism of bipolar depression.

In Tables 13, and 19, there are listings of candidate genes and specific single nucleotide polymorphisms that may be critical for the identification and stratification of a patient population diagnosed with depression based upon genotype. Current pathways that may have involvement in the therapeutic benefit of depression include glutaminergic, serontonergic, dopaminergic, adrenergic, cholinergic, purinergic, GABAergic, melatonin, peptide protein processing, opiates, oxytocin, neuropeptide Y, calcitonin/calcitonin gene related peptide, tachykinin, corticotropin releasing factor, vasopressin, calcium or potassium channels, prostaglandin, testosterone, oxygen stress, second messenger cascade, folate metabolism pathways that are listed in Tables 2, 7, 13, and 19. One skilled in the art would be able to identify these pathway specific gene or genes that may be involved in the manifestation of depression, are likely candidate targets for novel therapeutic approaches, or are involved in mediating patient population differences in drug response to therapies for depression.

In Table 28 below is a list of the available candidate therapeutic alternatives available or in development for depression. There are listed by therapeutic approach are defined and listed in Table 2. The listed candidate therapeutic interventions response in patients with depression may be affected by polymorphims in genes as described above.

**Example 13**

Epilepsy

I. *Description of Epilepsy*

Epilepsy is a neurological disorder affecting an estimated 1.8 million Americans with estimated direct and indirect costs of illness to be approximately $3 billion dollars. Epilepsy is characterized by the behavioral consequences of recurrent, spontaneous, transient paroxysms of abnormal brain activity. An epileptic attack or seizure may result in impaired consciousness, involuntary movements, autonomic disturbances, psychic or sensory disturbances. The fundamental etiology of epilepsy is thought to occur within the cerebral cortex or limbic cortex (hippocampus). Chronic epilepsy is the syndrome in which recurrent neuronal paroxysms that underlie ictal events are transient expression of more permanently physiological disordered cortex. In ascertaining the location and the diagnosis of epilepsy, one can determine patterns of uncoordinated cortex by examination of ictal and interictal EEG recordings. Interictal recordings of epilepsy patients have an appearance of brief discharges that can be recorded from the scalp. There is a noticeable spike-wave complex that is evident and is characterized by sharp negative transients followed by a slower wave. The EEG spike-wave complex reflects a summation of highly synchronized abnormal neuronal membrane potentials that upon inspection appear as large paroxsymal depolarization shifts followed by prolonged after depolarizations.

Epilepsy can be divided into the following categories based upon etiology: 1) primary epilepsy which is an intrinsic, nonprogressive, hereditary group of cerebral disturbances, 2) secondary epilepsy which is symptomatic of some know pathologic processes affecting the brain, and 3) reactive seizures which are characterized by natural reaction to physiologic stress or transient ischemic injury.

Epilepsy can be categorized into the following categories: partial seizures, generalized seizures, and seizures of unknown origin. Partial seizures are initiated (uni- or bilaterally) in discrete focal areas in the cortex and remain focal lesions. Generalized seizures begin either uni- or bilaterally and spread throughout the cortical tissue. In either case the mechanism of epileptogenic activity is to date unknown. However, there is evidence suggesting the etiology of epilepsy.

Partial seizures can be further subcategorized into: 1) simple partial seizure disorders, consciousness not impaired (with motor signs or symptoms, with somatosensory or special sensory symptoms (e.g. simple hallucinations, such as tingling, light flashes, buzzing), with autonomic signs and symptoms (e.g. epigastric sensation, pallor, sweating, flushing, piloerection and pupilary dilation), with psychic symptoms (e.g. disturbances of higher cerebral function (déjà vu, fear, distortion of time perception)); or 2) complex partial seizure disorders (simple partial onset following impairment of consciousness, impairment of consciousness at onset); or 3) partial seizures evolving to generalized tonic clonic seizures (simple partial seizures evolving to generalized seizures, complex partial seizures evolving to generalized seizures, and simple partial seizures evolving to complex partial seizures and further evolving to generalized seizures). A key feature of partial epilepsy is auras. These somatosensory or special sensory symptoms manifest as sensations described above and precede the seizure. There are cases whereby pharmacotherapy reduces the frequency and severity of partial seizures but may have little to no effect on aura sensation in partial epilepsy patients.

Generalized seizures are divided into 1) nonconvulsive seizures (absence seizures, atypical seizures, myoclonic seizures, or atonic seizures), or 2) convulsive seizures (tonic-clonic seizures, tonic seizures, or clonic seizures). Other seizure disorders that do not fit into the above categories are some cases of neonatal and infantile seizures.

There are other factors that one must consider when diagnosing seizure disorders. A generalized seizure may be the result sleep deprivation, alcohol or sedative drug withdrawal, use of convulsant drugs, fever, or acute head trauma. Furthermore, reversible toxic, infectious, or metabolic processes may induce recurrent generalized convulsions. Infantile febrile convulsions are an example of infancy and early childhood seizures that may or may not be indicative of a future epilepsy diagnosis.

Acquired epilepsy may be the result of congenital lesions, head trauma, infectious processes, brain tumors, cerebrovascular disease, systemic toxic and metabolic disturbances, hippocampal sclerosis, and miscellaneous disorders (collagen vascular disease, blood dyscrasias, cerebral gray matter degenerating diseases (allergic encephalopathy), presenile or senile dementias).

Epilepsy may be confused with clinical signs and symptoms of syncope, migraine, or pseudoseizures (nonepileptic psychogenic seizures). Usually, video/EEG monitoring of the patient during ictal and interictal periods allows trained personnel to distinguish epilepsy from these other clinical presentations.

*II. Current Medical Management of Epilepsy*

For the majority of patients, epileptic seizures can be controlled with antiepileptic drug therapy (in many cases, monotherapy) and may be withdrawn if the patient is seizure free for an extended period, usually 2 years. Some patients do not become free of seizures, despite therapy compliance. Persistent epilepsy, aside from deleterious effects on health, has psychosocial, behavioral, and cognitive consequences, which often impose financial burdens to patients, their loved ones, and society.

Based upon accurate diagnosis of the seizure type and seizure-associated physiology, appropriate therapy to reduce seizure frequency, severity, and epilepsy-associated behaviors can be identified. Diagnosis of epilepsy involves both identification of the epileptic syndrome and the type of seizure. Syndromes are identified based upon age of onset, EEG recording analysis, location of the epileptic region or site of epileptogenesis, type of seizure. The drugs available for medical management of epilepsy are divided by their use in the clinic; common forms of epilepsy are treated differently that partial or secondarily generalized tonic-clonic seizures disorders.

The current pharmacotherapy has three main mechanisms of action: 1) reduction of sustained repetitive firing of a neuron by promoting the inactivation state of voltage-activated Na+ channels; 2) enhanced GABAergic mediated presynaptic or postsynaptic inhibition of neural transmission; or 3) limiting the activation of specific voltage-activated Ca++ channels (T current). Following these general mechanism of action, current anticonvulsant drugs act by 1) prolonging the inactivation of the Na+ channels thereby reducing the ability of neurons to fire at high frequencies, 2) affecting GABAergic neurotransmission by reducing the metabolism of GABA, acting at the GABA receptor, enhancing the $Cl^-$ influx in response to GABA postsynaptically, or promoting presynaptic GABA release, or 3) reducing the flow Ca++ T-type calcium channels reducing the pacemaker current that underlies the thalmic rhythm in spikes and waves in generalized absence seizures.

There are generally accepted first- and second-line drugs for each of the types of epilepsies and associated syndromes. For partial seizures they are carbamazepine and phenytoin (first-line) and gabapentin, lamotrigine, phenobarbital, primidone, tagabine, topiramate and valproic acid (second-line). For generalized seizures they are: absence seizures ethosuximide and valproic acid (first-line); lamotrigine (second-line); myoclonic seizures, valproic acid (first-line), acetazolamide, clonazepam, lamotrigine, or primidone (second-line)); tonic-clonic seizures valproic acid, carbamazepine, phenytoin (first-line), lamotrigine,

phenobarbital, primidone (second-line); absence epilepsy with onset in childhood ethosuximide (first-line), valproic acid, lamotrigine (second-line); absence seizures with onset in adolescence valproic acid (first-line), ethosuximide, lamotrigine (second-line)); juvenile myoclonic epilepsy valproic acid (first-line), acetazolamide, clonazepam, primidone, lamotrigine (second-line); infantile spasms (West's syndrome corticotropin (first-line), clonazepam, valproic acid)); Lennox-Gastaut syndrome valproic acid, lamotrigine (first-line), carbamazepine (second-line).

Because there is greater risk for refractory epilepsy in partial epilepsy patients, there has been greater demand for the development of novel treatment alternatives. Since 1993 and as stated above, the introduction of lamotrigine, topiramate, tiagabine, and gabapentin have changed the medical management of partial epilepsy. Although carbamazepine and phenytoin remain the mainstay therapies, these additions to the antiepileptic arsenal have provided therapeutic alternatives to this subset population of epilepsy patients.

In addition to AEDs, refractory epilepsy may benefit from surgical therapy to remove the site of epileptogenesis or implantation of a device to stimulate the vagus nerve. Surgical removal of cortical tissue can be successful therapy in up to two thirds of certain selected epilepsy and can reduce the seizure frequency and severity in others. However, surgical therapy of refractory epilepsy is underused, and is often a delayed procedure. It has been estimated that there are approximately 50,000 epilepsy patients that could benefit from resective surgery, however, there are only an estimated 1,500 surgeries performed each year. Potential reasons for the profound difference in the potential number of surgical candidates and the number of procedures include: limited number of surgical teams performing the resective surgery; failure of primary physicians to identify potential candidates and to refer them to surgical centers; reluctance of third party payers to provide coverage for the costly presurgical diagnostic testing and procedures; and further, a reluctance on the part of the patient to voluntarily elect removal of cortical tissue.

Vagal nerve stimulation for the treatment of some patients with epilepsy has proven to be safe and well tolerated. A device is implanted in the upper quadrant that can be programmed to directly stimulate the vagal nerve. Stimulation of this autonomic nerve has lead to a documented 25% reduction of seizure frequency in refractory patients. The device does not appear to have similar efficacy when implanted in a partial epilepsy patient population. The use of the surgically implanted device has recently only been approved in the U.S. (June, 1997) for patients over 12 years of age with known refractory partial epilepsy. Transient hoarseness is a frequent side-effect of this device as a result of over-stimulation of the vagal nerve.

*III. Limitations of Current Therapies for Epilepsy*

The limitations of current medical management of epilepsy are 1) partial response to therapy or refractory epilepsy, 2) undesired side effects, 3) continuing medical management of refractory or partial response in epilepsy patients, and 4) noncompliance.

Partial Response to Therapy and Refractory Epilepsy as a Therapeutic Limitation

Approximately 80% of patients with epilepsy are medically managed with current pharmacotherapy. In the remaining 20%, epileptic seizure frequency and severity are refractory to currently available medications. Medical personnel are left with attempting combination therapy of available anti-convulsive therapy. Standard regimens of multiple anticonvulsant therapy are fraught with greater tendency towards unwanted side effects. Interestingly, 20% of the primary generalized epilepsy patients and 35% of partial epilepsy patients are refractory. A poor response to anti-epileptic therapy may be result of many different therapeutic or diagnostic causes. Since the focus of therapeutic management of refractory epilepsy is combination antiepileptic drug therapy, the balance of beneficial therapy and the patient's intolerance of the adverse effects of the AEDs must be appropriately monitored.

Undesired Side Effects or Toxicities as a Therapeutic Limitation

All of the anti-epilepsy agents or compounds have undesired side effects. For example, nausea, dizziness, diplopia, ataxia, sedation, impaired mentation, hyperactivity, folic acid deficiency, leukopenia, elevated serum alkaline phosphatase levels, pruritis, blood dyscrasias, hirsutism, gingival hyperplasia, coarsening features, weight gain, and alopecia have been described for various anticonvulsant therapies.

Individuals with epilepsy have an increased rate of mortality as compared to the general population. Mortality is associated with treatment and with seizures and may include one or more of the following: trauma, burns, and drowning, habitual seizures with cardiopulmonary disease, severe aspiration, food bolus, and sudden unexplained death. Sudden unexplained death in epilepsy patients (SUDEP) has been reported as high as 1 in 270 patients that are refractory to antiepilepsy drugs, and is a statistic that does not include suicides.

Additional concern of combination therapy besides increased propensity to experience undesirable side effects is the effect of metabolic rates and blood levels of the combinations. There is ample literature on the effect one antiepileptic agent

has on another, for example carbamazepine decreases the blood levels of clanzepam, ethosuximide, methsuximide, primidone, tiagabine, topiramate, and valproic acid while increasing phenobarbital blood levels. Clonazepam decreases the blood levels of carbamazepine while decreasing primidone blood levels.

Continuous Medical Management as a Therapeutic Limitation

Antiepileptic drug (AED) therapy of epilepsy requires continuous medical monitoring. Factors involving lifestyle may trigger seizures in a patient diagnosed with epilepsy who have seemingly medically managed disease. For example, emotional stress, sleep deprivation, menstrual cycle, flickering lights and other sensory stimuli, alcohol use or withdrawal, or comorbidities (i.e. infections) may exacerbate seizures.

Noncompliance as a Limitation of Current Therapies

Noncompliance or partial compliance is a major concern in both monotherapy or combination therapy. Many patients who are in what appears to be remission , tend to noncompliance of their prescribed therapy. Determining plasma levels of the drug or drugs can monitor compliance, but this places an added burden on the patient and family members. Noncompliance can result from additional factors: missed medication, failure to refill the medication, a complicated dosing regimen, problems with memory or vision, postictal confusion, denial of medical condition, fear of teratogenic effects of the drug or drugs during pregnancy, concerns about the effects (both short and long-term) of the medication, and inability to afford the medication.

Clearly, for some patients, refined therapeutic management of seizure frequency and severity would have benefits above and beyond the clinical setting. Without many therapeutic alternatives to refine combination antiepileptic agent regimens, epilepsy poses a continued impact on health-related quality of life for each patient.

*IV. Impact of Pharmacogenomics on Drug Development for Epilepsy*

Genetic mechanisms of epilepsy have recently been described. However, the clinical genetics of seizure disorders has been a relatively slowly progressing field. Molecular genetic approaches have been useful to identify genes or gene clusters involved in linkage analysis.

Genetic polymorphism analysis and effects of antiepileptic drug therapy was recently described for the cytochrome P450 2C9 and 2C19 genes and these variance differences on the metabolic rates of phenytoin. The polymorphisms considered in

this study were the arg144cys and the ile359leu of the CYP2C9 gene and the *1, *2, and *3 polymorphisms of CYP2C19. In this study of 134 Japanese patients, the mean maximal metabolic rates of phenytoin were 42% lower in individuals having the ile359leu genotype. From this analysis, the authors conclude that patients with the ile359leu genotype may not tolerate higher daily doses of phenytoin and may require genetic identification prior to implementation of medical strategies.

The evidence that a variance in a gene involved in a pathway that affects antiepilepsy drug response, indicates and supports the expectation that there is a likelihood that other genes have similar qualities to various degrees. As drug research and development proceeds to identify more lead candidate therapeutic interventions for epilepsy, there is possible utility in stratifying patients based upon their genotype for these yet to be correlated variances. Further, as described in the Detailed Description, methods for the identification of candidate genes and gene pathways, stratification, clinical trial design, and implementation of genotyping for appropriate medical management of a given disease is easily translated for epilepsy. As described below in section V. below there are likely gene pathways as are those that are outlined in the gene pathway Table 2 and in the gene pathway and indication matrix Table 7.

Optimization of GABAergic or ion channel modulation mediated therapy of epilepsy further demonstrates the utility of selection of a potential epilepsy patient that has a predisposing genotype in which selective AED or agents are more effective and or are safer. In considering an optimization protocol, one could potentially predetermine variance or variances within the GABAergic receptor, ion channel or ion channel mediated mechanisms of neurotransmission, or GABAergic receptor mediated intracellular mechanism of action that is preeminently responsible for drug response. By embarking on the previously described gene pathway approach, it is technically feasible to determine the relevant genes within such a targeted drug development program for epilepsy.

A sample of therapies approved or in development for preventing or treating the progression of symptoms of epilepy currently known in the art is shown in table 29. In this table, the candidate therapeutics were sorted and listed by mechanism of action. Further, the product name, the pharmacologic mechanism of action, chemical name (if specified), and the indication is listed as well.

## V. *Mechanism of Action Hypotheses for Novel Therapies for Epilepsy: Utility of Genotyping*

Further studies have demonstrated that there is a genetic component to epilepsy. These genetic factors may predispose by an individual to epilepsy by

inheriting one or more of the following 1) low threshold for aberrant seizure activity; 2) traits that underlie certain specific primary epilepsy disorders; and 3) a disease of the CNS in which there are associated structural disturbances that produce seizures. As described above there is an urgent need for the discovery and development of therapeutic alternatives for the medical management of epilepsy. Recent research and development programs have included the following exemplary hypothesis testing programs. In a first hypothesis, glutamate neurotransmitter pathway has been implicated in aberrant excitatory neurotransmission. Glutamate and aspartate are ligands for the N-methyl-D-aspartate receptors and ionophore receptors (AMPA and Glu 1-4). Research efforts have been dedicated to eliciting glutaminergic specific antagonists that may be productive inhibitors of aberrent exicitatory neural signals or may be effective to attentuate neural modulatory mechanisms that are defective in epileptogenic tissue.

Another hypothesis includes the glycinergic pathway. Because glycine is an additional excitatory neurotransmitter, efforts to identify glycinergic specific ligands that may be of therapeutic benefit to prevent, reduce, or ablate seizure activity in cortical tissue. A third hypothesis is the histamine receptor ligands and tachykinin receptor ligands may be useful for neuromodulation of excitatory neurotransmission.

Further, there may be genes within pathways that are either involved in metabolism of neurotransmitters or are involved in metabolism of various drugs or compounds. In Tables 1-6, 12-17 and 18-23, there are listings of candidate genes and specific single nucleotide polymorphisms that may be critical for the identification and stratification of a patient population diagnosed with epilepsy based upon genotype. Current pathways that may have involvement in the therapeutic benefit of epilepsy include glutaminergic, serontonergic, dopaminergic, adrenergic, cholinergic, purinergic, GABAergic, glycinergic, taurine, oxytocin, vasopressin, calcium, potassium, or sodium channels, mitochondrial maintenance, protein maturation and degradation, and second messenger cascade gene pathways that are listed in Tables 1-6, 12-17, 18-23. One skilled in the art would be able to identify these pathway specific gene or genes that may be involved in the manifestation of epilepsy, are likely candidate targets for novel therapeutic approaches, or are involved in mediating patient population differences in drug response to therapies for epilepsy.

Based upon these varying hypotheses there are many products in development for epilepsy. Table 29 below lists current therapies that have not yet received U.S. marketing approval. The listed candidate therapeutic interventions

response in patients with epilepsy may be affected by polymorphims in genes as described above.

**Example 14**

Migraine

*I. Description of Migraine*

Migraine is a neurological syndrome that has multiple, complex manifestations. Migraine with auras, unilateral throbbing, and associated nausea is the basic clinical symptomatic presentation. The premonitory phase may be up to 24 hours and may be associated with auras or alterations of mood, appetite, visual, sensory, or motor functions. Migraine headache is a unilateral throbbing that is associated with photophobia, hypacusis, polyuria, and diarrhea.

There are many clinical subtypes of migraine. Broadly, these subtypes can be distinguished by the presence or absence of auras. Migraines without auras are defined as the classic type. Migraines with auras can be further classified as 1) migraine with typical auras, 2) migraine with prolonged auras, 3) familial hemiplegic migraine, 4) basilar migraine, 5) migraine without headache, and 6) migraine with acute-onset aura. Additional migraine types include ophthalmologic migraine and retinal migraine.

II. *Current therapies for Migraine*

Migraine medical therapy depends on the acute or prophylactic nature of the therapy and whether the migraine is diagnosed as mild, moderate, or severe. Many patients will take a step approach to each separate migraine attack, starting with weakly acting agents and progressing to more potent drugs. For patients with severe migraine, therapy includes prophylactic management.

Therapy for Acute Migraine

Mild migraine is a headache that may accompanied by nausea, is unilaterally throbbing, and can be treated by nonprescription analgesics. Patients infrequently consult a neurologist for care of mild migraines because the level of impairment imparted by the headache portion is not debilitating and is relatively short lived. Mild migraine is thus treated with aspirin, acetominophen, ibuprophen, indomethacin, naproxen sulfate, and in some cases antiemetic drugs (diphenhyrdamine, prochlorperazine, promethazine, and metchlorpramide).

Moderate migraine is generally characterized by similar symptoms of mild migraine, however the frequency and or severity are increased. Patients with

moderate migraine are generally not relieved with non-narcotic analgesics, and require medications that combine aspirin or acetominophen with a mild sedative or α and β adrenergic receptor mediated vasoconstriction.

Severe migraine is characterized by similar symptoms as mild and moderate migraine. However, the severity and frequency of headache is debilitating. Patients seek relief from the headache pain in the acute stage and frequently require prophylactic maintenance therapy. The drugs used for the therapy of acute migraine are members of the ergot alkaloid family or sumitriptan.

The ergot alkaloids are partial agonists and antagonists for a variety of receptor types; serotoninergic, adrenergic, dopaminergic, muscarinic, and GABAergic. Synthetic products with similar chemical structures to ergotamine predominantly are agonists at the serotonin subtype 1D or 1B. Both of these two subtypes act by inhibiting adenylyl cyclase activity in cortical neurons. Ergotamine is also a vasoconstrictor; this activity is thought to occur through activation of the α1 adrenergic receptor system. Ergotamine is metabolized by undefined pathways and metabolites are excreted primarily in the bile. The bioavailability of ergotamine is approximately 1% due the potent first pass effect after parenteral delivery of the drug and erratic absorption between individuals.

Sumitripan is another drug used for the acute, severe migraine attacks. Sumitripan is a serotonin 1B, 1D selective receptor agonist. Because these receptor subtypes are auto receptors, activation of 5HT1B and 5HT1D receptors can act by controlling the release of the serotonin and other neurotransmitter release. Sumitriptan may also be efficacious in the treatment of migraine because it may block proinflammatory receptors at the level of nerve terminal in the perivascular space.

Other drugs used as adjunct therapy for acute, severe migraine attacks are corticosteroids and opioid analgesics. Due to their addictive qualities, opioid or narcotic analgesics are limited to acute, infrequent attacks.

Prophylactic Therapy of Migraine

There are currently six classes of standard treatments for the prophylactic therapy of migraine. They are 1) tricyclic antidepressants (amytriptylline), 2) 5HT antagonists (methylsergide), 3) β adrenergic receptor antagonists (propanolol, timolol, atenolol, metropolol, nadolol), 4) monoamine oxidase inhibitors (depranil), 5) calcium channel blockers (verapamil, flunarizine), and 6) anticonvulsants (divalproex sodium, chlorpromazine). The criteria for the selection of prophylactic therapy are 1) patient has 6 or more headaches each month, 2) symptomatic medications are contraindicated or ineffective, 3) medication is necessary more than

twice each week, and 4) there is an expressed need on the part of the patient to prevent infrequent attacks, e.g. hemiplegic migraine, those headaches producing profound disruption, or those associated with a risk of stroke. The ultimate choice of the prophylactic medication is based upon the measured effect on the type of migraine and the patient's willingness to withstand the associated side effects.

*III. Limitations of Current Therapies for Migraine*

The current therapy of migraine includes management of acute attacks of the mild, moderate and severe categories. Therapies of severe migraine further include prophylactic therapies. Regardless of the acute or prophylactic nature of the therapy, there are both efficacy and toxicity limitations in which migraine remains problematic for medical management.

Toxicity or Undesired side effects of Acute Migraine Therapy

Ergotamine and its derivatives are useful drugs for the management of acute severe migraine attacks, however there are side effects associated with administration of the drug. Ergotamine is an activator of the CNS emetic centers, and nausea and vomiting are a frequent side effect that can be confused with a manifestation of the migraine attack. Other undesirable side effects are weakness of the legs, muscle pains, numbness and tingling of toes, and transient tachy- or bradycardia.

A known side effect of sumitriptan is coronary vasospasm and it is thus contraindicated in patients with ischemic heart disease or Prinzmetal's angina.

Limitations of Prophylactic Migraine Therapy

Although prophylactic therapy for migraine can reduce the frequency and intensity of the migraine attack, there are patients that achieve dramatic improvement and there are those that achieve only a 50% reduction, indicating a limited efficacy and benefit for a significant population subgroup. In those patients, the severity and intensity must be significant to require daily prophylactic medication.

Of the six categories of prophylactic agents all have associated side effects that may or may not be tolerable to each individual patient. They are 1) tricyclic antidepressants: sedation, dry mouth, weight gain, tremor, cardiac arrythmias, aggravation of angle-closure glaucoma, and difficulty in urinating; 2) 5HT antagonist: weight gain, muscle cramps, vasoconstriction, and retroperitoneal pleuroperitoneal and subendocaridal fibrosis; 3) β adrenergic receptor antagonists: aggravation of asthma, bradycardia, hypotension, fatigue, depression, masking the

symtpoms of diabetes mellitus; 4) monoamine oxidase inhibitors: orthostatic hypotension, insomnia, and nausea; 5) calcium channel blockers: are not frequently used, however are associated with constipationa and orthostatic hypotension; and 6) anticonvulsants: nausea, fatigue, weight gain, alopecia, tremor, liver dysfunction, and neural tube defects in developing embryos.

The least desired effect of prolonged prophylactic therapy of migraine is the associated increased frequency of headaches. Headaches, not of the migraine type, can occur daily and are related to rebound withdrawal from frequent use of the acute antimigraine medication. Patients experiencing this type of headache pattern are said to have transformed migraine and often experience episodic migraine attacks superimposed on their daily headaches. Ergotamines are frequently associated with chronic daily headaches, as are the triptans. Unfortunately, patients experiencing daily headaches are less likely to respond to acute therapy or any other preventative medications. Withdrawal of other migraine medications further render the patient more susceptible to daily headaches. Therefore, it is beneficial to prevent transformed migraine and chronic daily headaches. Drugs known to be associated with transformed migraine are generally limited to occasional use in patients that have greater than two migraines each month. It is additionally recommended for patients that experience more frequent headaches requiring over-the-counter or prescription medications be put on a rotating schedule.

*IV. Impact of Pharmacogenomics on Drug Development for Migraine*

As described above, there is evidence to suggest that there are efficacy and safety different responses to drug therapy within the migraine patient population. Although not all of these responses may be attributable to genotypic differences, it is expected that if stratification based upon genotype were performed, a reasonable correlation between drug response and genotype may become obvious. As described below, there are gene pathways that are involved with current drug therapy and those that may be potentially involved in the future. As described in the Detailed Description, methods for the identification of candidate genes and gene pathways, stratification, clinical trial design, and implementation of genotyping for appropriate medical management of a given disease is easily translated for migraine and patients diagnosed with migraine. As described below in section V. below there are likely gene pathways as are those that are outlined in the gene pathway Table 2 and matrix Table 7.

Optimization of serontonergic, nonsteroidal antiinflammatory, or cerebral vasoconstrictor mediated mechanism of therapy of migraine further demonstrates the

utility of selection of a potential migraine patient that has a predisposing genotype in which selective anitmigraine or agents may be more effective and or have an more desirable safety profile. In considering an optimization protocol, one could potentially predetermine variance or variances within the serotonergic receptor pathway, nonsteroidal aninflammatory pathway, or serotonergic receptor or nonsteroidal antiinflammatory mediated intracellular mechanism of action that is preeminently responsible for antimigraine drug response. By embarking on the previously described gene pathway approach, it is technically feasible to determine the relevant genes within such a targeted drug development program for migraine.

A sample of therapies approved or in development for preventing or treating the progression of migraine currently known in the art is shown in Table 31. In this table, the candidate therapeutics were sorted and listed by mechanism of action. Further, the product name, the pharmacologic mechanism of action, chemical name (if specified), and the indication is listed as well.

Pharmacogenomics studies for these drugs, as well as other agents, drugs, compounds or candidate therapeutic interventions, could be performed by identifying genes that are involved in the function of a drug including, but not limited to is absorption, distribution metabolism, or elimination , the interaction of the drug with its target as well as potential alternative targets, the response of the cell to the binding of a drug to a target, the metabolism (including synthesis, biodistribution or elimination) of natural compounds which may alter the activity of the drug by complementary, competitive or allosteric mechanisms that potentiate or limit the effect of the drug, and genes involved in the etiology of the disease that alter its response to a particular class of therapeutic agents. It will be recognized to those skilled in the art that this broadly includes proteins involved in pharmacokinetics as well as genes involved in pharmacodynamics. This also includes genes that encode proteins homologous to the proteins believed to carry out the above functions, which are also worth evaluation as they may carry out similar functions. Together the foregoing proteins constitute the candidate genes for affecting response of a patient to the therapeutic intervention. Using the methods described above, variances in these genes can be identified, and research and clinical studies can be performed to establish an association between a drug response or toxicity and specific variances.

*V. Description of Mechanism of Action Hypotheses for Future Migraine Drug Development*

The pathogenesis of migraine includes the following theories: vascular, depression of cortical electrical activity, serotonergic abnormalities, alteration of

neurotransmitter modulation, and modulation of neuroendocrine mechanisms. These are described briefly below.

The vascular theory of migraine posits that there is abnormal cerebral blood flow and it apparently plays a pivotal role in the clinical symptoms of migraine. Studies have shown that a decrease in cerebral blood flow during an aura and an increase in blood flow during headache occur in migraine patients. This theory is somewhat substantiated indirectly by the pharmacologic action of therapies for acute migraine and prophylaxis, as previously described.

There have been additional studies that point to a mechanism of spreading depression of cortical electrical activity and a concurrent alteration of blood flow. This theory suggests that focal reduction of electrical activity and concurrent increase in blood flow occurs focally and spreads across the hemisphere at a rate of 2-3 mm each minute. This spreading hypothesis has been refined to a description of migraine as an evolving process in the cerebral cortex that occurs secondarily to decreased cortical function, decreased coritcal metabolism, and or vasoconstriction of cortical arterioles.

Many studies have addressed the effect of serotonergic mechanism of the pathogenesis of migraine. These studies used the following premises: 1) there have been reports of decreased concentrations of serotonin in platelets and plasma, 2) increased levels of serotonin and serotonergic metabolites in urine, 3) lastly, migraine may be precipitated by abnormal release of biogenic amines, a theory borne out of the fact that reserpine and fenfluramine can precipitate a migraine attack.

Other theories propose that alteration s of neurotransmitter systems e.g. nitric oxide, glutamate, and opioid receptors may be part of the pathogenesis of migraine. Further, Some studies have included anatomical differences in the raphe system and within the cerebral vasculature as well as alterations of the autonomic nervous system.

Therapy of migraine is dependent on the appropriate diagnosis, as well as the type, frequency, and severity of the throbbing headache. Upon diagnosis, patient education to identify and avoid trigger factors is a critical first step in all patients.

Trigger factors may include but are not exclusive to alcohol (red wine), foods (chocolate, certain cheeses), irregular sleep patterns, and acute changes in stress levels. Triggers may also come from environmental factors, such as time-zone shifts, high altitudes, or barometric changes. In women, menstrual cycles may trigger a migraine attack. These trigger factors suggest that there are complicating factors to include in any pathophysiologic hypothesis of migraine, and that these hypotheses may include neuroendocrine, endocrine, and other metabolic factors.

Further, there may be genes within pathways that are either involved in metabolism of neurotransmitters or are involved in metabolism of various drugs or compounds. In Tables -6, 12-17 and 18-23, there are listings of candidate genes and specific single nucleotide polymorphisms that may be critical for the identification and stratification of a patient population diagnosed with epilepsy based upon genotype. Current pathways that may have involvement in the therapeutic benefit of migraine include glutaminergic, serontonergic, dopaminergic, adrenergic, cholinergic, GABAergic, nitric oxide, peptide hormone processing, opiates, tachykinin, bradykinin, corticotropin releasing hormone, calcitonin/calcitonin gene related peptide, calcium channel, hemostasis, and second messenger cascade gene pathways that are listed in Tables -6, 12-17 and 18-23. One skilled in the art would be able to identify these pathway specific gene or genes that may be involved in the manifestation of migraine, are likely candidate targets for novel therapeutic approaches, or are involved in mediating patient population differences in drug response to therapies for migraine.

Based upon these varying hypotheses as stated above, there are many products in development for migraine. Table 31 below lists current therapies that have not yet received U.S. marketing approval.

**Example 15**

Psychosis

Psychosis is a general term for major mental disorder characterized by loss of contact with reality, often manifested by disordered thought, delusions or hallucinations. Psychosis can be part of several distinct psychiatric diseases, including schizophrenia, manic-depressive disease, severe depression with psychotic features, organic psychotic disorders, as well as in alcohol or drug intoxication and acute idiopathic psychotic illnesses. The most common of these is schizophrenia. The antipsychotic drugs are also used to treat non-psychiatric conditions such as, for example, nausea and vomiting, movement disorders associated with neurodenerative diseases such as Huntington's disease and Tourette's syndrome, pruritis and chronic hiccough. Example 11 focuses predominantly on schizophrenia, however similar analysis, in terms of the relevant pathways, genes, polymorphisms and analytical methods for establishing relationships between polymorphisms and drug responses, would obtain in all the other diseases treated with antipsychotic drugs. Criteria for the diagnosis of schizophrenia and other psychoses, as well as diagnostic criteria for the other disorders treated with antipsychotics, are well established. (Diagnostic and

Statistical Manual of Mental Disorders, 4th ed., American Psychiatric Association Press, Washington, D.C., 1994.)

II. Current Medical Management of Schizophrenia

Over 15 drugs are approved for treatment of psychosis in the US. They include the so-called conventional or typical antipsychotic drugs and the more recently introduced atypical antipsychotic drugs.The former class includes phenothiazines (e.g. chlorpromazine, the first antipsychotic to be widely used), thioxanthenes (e.g. thiothixene), butyrophenones (e.g. haloperidol, one of the most useful conventional antipsychotics) and other heterocyclic compounds. The atypical antipsychotics include compounds such as clozaril (the first, and best studied member of the class), risperidone, olanzapine, quetiapine, ziprasidone and iloperidone. Some drugs, such as loxapine, have pharmacology intermediate between that of the typical and atypical drugs.

The typical antipsychotics are believed to act predominantly by antagonizing dopamine receptors, particularly D2-dopamine receptors.. These medications can be effective in reducing the positive symptoms of schizophrenia (hallucinations, delusions) but are generally not effective at alleviating the negative symptoms (withdrawal, flat affect, anhedonia, lack of will), nor do they generally result in improved cognitive function. In fact, negative symptoms and cognitive function may worsen on typical antipsychotics. Typical antipsychotics exhibit dose dependent efficacy, and the optimal dose for a given patient must be determined empirically by gradually increasing the dose until adequate control of symptoms is achieved (without unacceptable side effects – see below). A therapeutic dose is usually reached within 2-3 weeks of initiating therapy.

The atypical antipsychotic drugs have replaced the typical agents as front line therapy for schizophrenia and other psychoses because they have a beneficial impact on the negative symptoms as well as the positive symptoms of schizophrenia, and because, based on recent research, they may also improve cognitive function. The atypical drugs affect a number of neurotransmitter systems, with modulation of serotonergic neurotransmission – particularly 5HT2C receptor antagonism, a prominent effect in addition to modulation of dopaminergic function.. The best studied of this class of drugs is clozapine, which binds dopamine receptors with low affinity, and also interacts with muscarinic, adrenergic, serotonergic, and histaminergic receptors. The table below depicts the relative receptor affinity (0-5

on a scale of 5, where 5 indicates a high affinity interaction) of a conventional drug (haloperidol) and an atypical drug (clozapine).

Relative Receptor Affinities of Haloperidol and Clozapine

| | Neurotransmitter Receptor Subtype | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | D1 | D2 | $5HT_{2A}$ | $5HT_{1A}$ | α1 | α2 | H1 | M1 |
| Haloperidol | +3 | +4 | +1 | 0 | +2 | 0 | 0 | 0 |
| Clozapine | +2 | +2 | +1 | +3 | +3 | +3 | +4 | +5 |

The effectiveness of the atypical antipsychotic drugs has revealed the inadequacy of a simplistic dopamine excess hypothesis of schizphrenia. The clinical effects of the atypical antipsychotic drugs are likely to reflect the summation of a complex set of interactions with a variety of neurotransmitter receptors. Interpatient differences in the function, levels or anatomical distribution of these different receptors are likely to account for a substantial fraction of interpatient variation in response to atypical antipsychotic drugs. Further, the function, levels and anatomical distribution of receptors is largely under genetic control, as is the associated biosynthetic, catabolic, recycling and signal transduction machinery. An understanding of the specific genetic variants that have major effects on drug efficacy would allow a far more sophisticated selection of appropriate therapy and dose than is possible currently.

*III. Limitations of Current Therapies*

The chief limitations of antipsychotic medicines are (i) conventional and atypical neuroleptic agents do not reduce the signs and symptoms of schizophrenia in all patients (an estimated one third to one quarter of psychotic patients are resistant to therapy); (ii) a wide range of serious adverse effects. Further, it is impossible to predict the response of any given patient, particularly the mix of drug effects on positive symptoms, negative symptoms, cognitive deficits and side effects. As a result, selection of therapy is at present completely empirical. This approach is costly, as (i) multiple physician visits may be required before an optimal dose of an effective agent is attained; (ii) even after determining an effective drug regiment, the long term effects of therapy in specific patients generally remain unknown, particularly with respect to side effects; (iii) these problems result in low rates of compliance with therapy. Hence there is a need for tools that would allow the prospective identification of patients likely to be responsive to - and free from short or long term side effects from - a particular drug.

Efficacy Limitations

The dilemma confronting psychiatrists and other clinicians selecting therapy for psychotic patients has been described by Baldessarini in Goodman and Gilman's The Pharmacological Basis of Therapeutics (9th edition) as follows: "No one drug or combination of drugs has a selective effect on a particular symptom complex in groups of psychotic patients; although individual patients may appear to do better with one agent than another, this can be determined only by trial and error". Thus, a clinician selecting therapy for a newly diagnosed psychotic patient, generally selects a compound with which he is comfortable, based on past experience. If that agent is not effective, or is producing a side effect, then a second agent is selected, again, entirely based on the physicians clinical judgement, and so on. This approach to optimization of pharmacotherapy has both medical and economic drawbacks. From the medical point of view, it does not always result in the selection of optimal treatment, with the attendant drawbacks in patient compliance. From an economic viewpoint the number of physician visits required to reach an effective dose of an effective drug are greater than necessary, and some patients may require hospitalization during the period when various drug regimens are being tested. There are other costs of using less than optimal therapy: (i) a patient might experience an improvement in cognitive symptoms on an optimal drug that would allow performance of a regular job; suboptimal therapy, even while adequately controlling positive symptoms, might not be sufficient to enable job performance. (ii) An optimal drug would minimize side effects, and thereby reduce physician visits, while also resulting in greater compliance. (Noncompliance is likely to ultimately lead to more hospitalization.)Determination of an optimal dose of an antipsychotic is another challenging aspect of therapy with these agents. Baldessarini (Goodman and Gilman, 9th ed.) writes: "Optimal dosage of antipsychotic drugs requires individualization to determine doses that are effective, well-tolerated, and accepted by a patient. Careful observation of the patients changing response is the best guide to dosage." As with selection of an optimal agent, discussed above, the determination of an optimal dose presently requires multiple physician visits. Clearly some fraction of interpatient variation in optimal dose level is likely due to genetic, and consequent biochemical differences between patients. Such differences may involve drug metabolizing enzymes or proteins that mediate pharmacodynamic effects. A list of such proteins is provided in Tables 1-6. Many typical antipsychotic drugs are metabolised by cytochrome P450 enzymes, with consequent wide interpatient variation in pharmacokinetic parameters. Further, many antipsychotic drugs are converted to active metabolites which can have

therapeutic effects or side effects. The metabolism of the tricyclic atypical drugs (clozapine, olanzapine, and quetiapine) occurs via N+-oxidation, N-glucuronidation, and phases 1 and 2 metabolism with final glucuronidation before renal excretion. The non-tricyclic atypical antipsychotic drugs (e.g. risperidone, sertindole and ziprasidone) have diverse chemical structures and there is less data on their metabolism, but it appears to include diverse phase 1 biotransformation reactions. As a rule, conventional antipsychotics are mainly effective against positive symptoms (hallucinations, delusions, illusions), while not significantly ameliorating negative symptoms (withdrawal and flat affect). They are also associated with a high incidence of adverse effects, particularly extrapyramidal symptoms (EPS) and tardive dyskinesia. Atypical antipsychotics constitute a significant improvement, in that they are at least as effective as conventional drugs against positive symptoms, they show at least some effectiveness against negative symptoms and, according to recent studies, they may also produce improvement in the cognitive deficits associated with schizophrenia (e.g. attention, executive function, short and long term memory), while causing substantially fewer extrapyramidal symptoms.

Toxicity Limitations

Unfortunately, conventional anti-psychotic drugs are uniformly associated with undesirable dose-dependent side effects. These include (but are not limited to) extrapyramidal effects, electrocardiogram abnormalities, sedation, weight gain, cognitive deficits, sexual or reproductive dysfunction, blood dyscrasias (particularly agranulocytosis associated with clozapine), , neuroleptic malignant syndrome (parkinsonism with catatonia), jaundice, skin reactions, epithelial keratopathy andseizures. Skin reactions include uticaria and dermititis and are usually associated with phenothiazines. Epithelial keratopathy and corneal opacities are associated with chlorpromazine therapy. In extreme cases these effects impair vision, but they tend to spontaneously disappear upon discontinuation of chlorpromazine.

The extrapyramidal side effects of conventional neuroleptics include dystonia (facial grimacing, torticollis, oculgyric crisis), akathesia (feeling of distress or discomfort leading to restlessness or constant movement), and parkinsonian syndrome (rigidity and tremor at rest, flat facial expression).

Tardive dyskinesia is a common side effect of long term usage of conventional neuroleptic drugs.Tardive dyskinesia is a syndrome of abnormal involuntary repetitive, painless movements. These movements vary in intensity over time, dependent on the level of arousal or emotional distress. Typically there are

quick choreiform (ticlike) movements of the face, eyelids (blinks or spasms), mouth (grimaces), tongue, extremities, or trunk. Increasing the dose of the conventional neuroleptic agent can reverse extrapyramidal effectsshort term, but at the cost of more severe dyskinesia long term. Not infrequently a clinician is compelled to change medication for a patient with adequately controlled schizophrenia because of dose related tardive dyskinesia or other extrapyramidal side effects..

Another important side effect of many antipsychotic drugs is QT wave prolongation, which has recently resulted in the withdrawal of an atypical antipsychotic compound. Cardiac conduction abnormalities associated with antipsychotic therapy have resulted in patient deaths, presumably as a consequence of ventricular tachycardias. The mechanism of the conduction abnormalities appears to involve drug binding to cardiac potassium channels and consequent interference with repolarization current. Sertindole, for example, is a new antipsychotic agent that binds with high affinity (3-14 nM, depending on conditions) to and antagonizes HERG, a cardiac potassium channel. The degree of interpatient variation in these effects is not well characterized. Genes likely to account for these differences encode potassium channels (which may also have some role in the central actions of these compounds), sodium channels and the genes associated with inherited forms of long QT wave syndrome (QT1, QT2, QT3, QT4, QT5 and QT6).

Yet another important side effect of antipsychotic drugs is weight gain which can lead to obesity.

IV. *Impact of Genotyping on Drug Development for Schizophrenia*

Most traditional neuroleptics have a narrow therapeutic-to-toxic index, and thus, the novel antipsychotics are the result of a search to substantially widen the distance between the dose that treats psychosis and the one that produces adverse effects. In vitro binding profiles have been created for the atypical antipsychotics that have been approved by the U.S. Food and Drug Administration (FDA)-clozapine, olanzapine, and risperidone and those that are under FDA review-quetiapine and sertindole. These profiles, which were compared with that of the typical neuroleptic haloperidol, provide guidance for predicting the adverse effects produced by these drugs. Most conventional antipsychotics have central nervous system effects, particularly extrapyramidal symptoms (EPS) and tardive dyskinesia, sedation, and dulling of cognition. Other adverse effects of the typical antipsychotics include the neuroleptic malignant syndrome, orthostatic hypotension, changes in liver function, anticholinergic and antiadrenergic side effects, sexual

dysfunction, and weight gain. The newer agents have a lower incidence of EPS and tardive dyskinesia, while weight gain and changes in blood pressure and liver function tests are adverse effects that have been associated with the use of the newer agents. The favorable side effect profile of these new antipsychotics is likely to make patients more willing to continue treatment, and thus these agents represent a step forward in the treatment of patients with severe, chronic mental illness.

This paper reviews the current literature describing the metabolism of both multi-receptor clozapine analogue atypical antipsychotic drugs (clozapine, olanzapine, and quetiapine) and serotonin-dopamine antagonist atypical antipsychotic drugs (risperidone, sertindole and ziprasidone), to highlight the significance of those data in the context of clinical practice. The former group of atypical antipsychotic drugs shares a similar tricyclic structural nucleus and are metabolized through three major categorical metabolic pathways--N+-oxidation, N-glucuronidation, and phases 1 and 2 biotransformation with final glucuronidation before renal excretion.

There have been reports of polymorphisms in key genes that affect neuroleptic activity in schizophrenic patients. For example, within the dopamine D4 receptor subtype, there are known tandem repeats in exon 3. In a recent study, schizophrenic patients on maintenance doses of chlorpromazine were stratified into two groups, one having 2 tandem base pair repeats and the other having 4 tandem base pair repeats. Thirty-four percent of group one patients and 62% of group two patients had a favorable response to chlorpromazine therapy during acute stage treatments. The presence of homogeneous four 48 base pair repeats in both alleles in exon 3 of the dopamine D4 receptor subtype thus appears to be associated with beneficial chlorpromazine response.

Recently, a study of the serotonin receptor subtype 6, polymorphism (T267T vs. C267T) in a group of patients refractory to clozapine therapy was reported. In this study, it appeared that the T267T genotype patients were more likely to respond to continued therapy that those patients with C267T genotype patients.
A recent report documented a correlation of the serotonin 5HTC2 receptor subtype biallelic polymorphism and neuroleptic efficacy. A significant number of schizophrenic patients homozygous for the allele C2 who responded unsatisfactorily to antipsychotic medication as compared to control.

Three polymorphisms in the serotonergic receptors, i.e. 5HT2A (T102C); 5HT2C (cys23ser); and 5HT2A (his452tyr) have reports of positive or negative correlation with efficacy of antipsychotic therapies. This disparity in the literature

will, in the future, be further examined in schizophrenic patient populations and correlation may be discovered.

The evidence that a variance in a gene involved in a pathway that affects neuroleptic drug response, indicates and supports the theory that there is a likelihood that other genes have similar qualities to various degrees. As drug research and development proceeds to identify more lead candidate therapeutic interventions for schizophrenia, there is possible utility in stratifying patients based upon their genotype for these yet to be correlated variances. Further, as described in the Detailed Description, methods for the identification of candidate genes and gene pathways, stratification, clinical trial design, and implementation of genotyping for appropriate medical management of a given disease is easily translated for schizophrenia. As described below in section V. below there are likely gene pathways as are those that are outlined in the gene pathway Table 2 and matrix table 7.

Identification of pathophysiologic relevant variance or variances and potential therapies affecting those allelic genotypes or haplotypes will speed the drug development. There is a need for therapies that are targeted to the disease and symptom management with limited or no undesirable side effects. Identification of a specific variance or variances within genes involved in the pathophysiologic manifestation of schizophrenia and specific genetic polymorphisms of these critical genes may assist the development of novel neuroleptic agents and the identification of those patients that may best benefit from therapy of these candidate therapeutic alternatives.

By identifying allelic variances or haplotypes in genes that indirectly affects efficacy, safety or both one could target specific secondary drug or agent therapeutic actions that affect the overall therapeutic action of conventional, atypical, or novel neuroleptic action.

A sample of therapies approved or in development for preventing or treating the progression of symptoms of schizophrenia currently known in the art is shown in Table 35. In this table, the candidate therapeutics were sorted and listed by mechanism of action. Further, the product name, the pharmacologic mechanism of action, chemical name (if specified), and the indication is listed as well.

V. *Mechanism of Action Hypotheses for Novel Therapies for Schizophrenia: Utility of Genotyping*

The underlying etiology of schizophrenia is not established, however there is compelling evidence that modulation of several neurotransmitter systems has an impact on the disease. As discussed above, conventional anti-psychotic drugs, effective in the management of schizophrenia, are dopamine antagonists, specifically D2-receptor antagonists, which block dopaminergic neurotransmission in the forebrain. Additionally, drugs such as mescaline and amphetamines, which are known to stimulate dopaminergic pathways, have been shown to induce psychotic symptoms. Evidence of dysfunctional serotonergic neurotransmission in schizophrenia includes evidence of altered serotonin receptor density, altered serotonin metabolism, and the evidence that serotonin receptors appear to be important targets for the atypical neuroleptics.

Based on current knowledge, there are three hypotheses that underscore the utility of polymorphic genotype analysis within the schizophrenic population. In the first, it could be considered that endogenous dopamine levels and consequential dopaminergic tone varies among schizophrenic patients, affecting response to receptor antagonists. These genetic DNA variations may affect brain neurotransmitter modulation of dopaminergic transmission and dopaminergic receptor mediated intracellular mechanisms among schizophrenic patients. In the second hypothesis, genetic DNA variations may affect the level of expression and brain distribution of dopamine receptors, imparting a variation in drug response among schizophrenia patients. Further, consideration of other endogenous neurotransmitters, i.e. serotonin, levels and consequential endogenous neurotransmitter tone varies among schizophrenic patients, affecting response to neurotransmitter receptor ligands or neurotransmitter receptor mediated intracellular mechanisms.

Further, there may be genes within pathways that are either involved in metabolism of neurotransmitters or are involved in metabolism of various drugs or compounds. In Tables1-6, 12-17, and 18-23, there are listings of candidate genes and specific single nucleotide polymorphisms that may be critical for the identification and stratification of a patient population diagnosed with epilepsy based upon genotype. Current pathways that may have involvement in the therapeutic benefit of schizophrenia include glutaminergic, serontonergic, dopaminergic, adrenergic, cholinergic, histaminergic, GABAergic, glycinergic, opiates, cholecystokinin, neurotensin, tachykinin, calcium channels, and second messenger cascade gene pathways that are listed in Tables 1-6, 12-17, and 18-23. One skilled in the art would be able to identify these pathway specific gene or genes that may be involved in the manifestation of schizophrenia, are likely candidate

targets for novel therapeutic approaches, or are involved in mediating patient population differences in drug response to therapies for schizophrenia.

**Example 16**

Effect of Pharmacokinetic parameters on Efficacy of Drugs and Candidate Therapeutic Interventions

The efficacy of a compound is determined by a combination of pharmacodynamic and pharmacokinetic effects. Both types of effect are under genetic control. In the present invention, the genetic determinants of efficacy are discussed in terms of variation in the genes that encode proteins responsible for absorption, distribution, metabolism, and excretion of compounds, i.e. pharmacokinetic parameters.

The pharmacokinetic parameters with potential effects on efficacy include absorption, distribution, metabolism, and excretion. These parameters affect efficacy broadly by controlling the availability of a compound at the site(s) of action. Interpatient variability in the availability of a compound can result in undertreatment or overtreatment, or in adverse reactions due to levels of a compound or its metabolite(s). Differences in the genes responsible for pharmacokinetic variation, therefore, can be a potential source of interpatient variability in drug response.

*Impact of Stratification Based Upon Genotype in Drug Development for Drugs, Compounds, or Candidate Therapeutic Interventions that may Efficacy*

Clozapine induced agranulocytosis has been associated in some reports with specific HLA haplotypes or with HSP70 variants. These reports suggest that a gene within the HLA region is associated with agranulocytosis in response to clozapine therapy. In a recent study, two ethnic groups were analyzed for genetic markers for agranulocytosis. Tumor necrosis factor microsatellites d3 and b4 were found in higher frequencies in patients that experience clozapine-induced agranulocytosis. These data, while they need to be confirmed by additional studies, are suggestive that tumor necrosis factor polymorphisms may also be associated with clozapine-induced agranulocytosis.

In this invention we provide additional genes and gene sequence variances that may account for variability in toxic reponses. The Detailed Description above demonstrates how identification of a candidate gene or genes (e.g. gene pathways), genetic stratification, clinical trial design, and diagnostic genotyping can lead to improved medical management of a disease and/or approval of a drug, or broader use of an already approved drug. Gene pathways including, but not limited to, those that are outlined in the gene pathway, Tables 1-6, preferably Table 3, are useful in

identifying the sources of interpatient variation in efficacy as well as in the adverse events summarized in the column headings of Table 8. Discussed in detail below are exemplary candidate genes for the analysis of pharmacokinetic variability in clinical development, using the methods described above.

Advantages of Inclusion of Pharmacogenetic Stratification in Clinical Development of Agents: Impact on Efficacy

As an example of identification of the primary gene and relevant polymorphic variance that directly affects efficacy, safety, or both one could select an gene pathway as described in the Detailed Description, and determine the effect of genetic polymorphism and therapy efficacy, safety, or both within that given pathway. For example, referring to Table 8, genes involved in absorption and distribution, phase I and phase II metabolism, and excretion the optimization of therapy of by an agent known to have an efficacious effect by determining whether the patient has a predisposing genotype in which the selected agents are more effective and or are more safe. In considering an optimization protocol, one could potentially predetermine the genotypic profile of these genes involved in the manifestation of the adverse effect, or those genes preeminently responsible for drug response. By embarking on the previously described gene pathway approach, it is technical feasibility to determine the relevant genes within such a targeted drug development program.

**Example 17**

Drug-Induced Toxicity: Blood Dyscrasias

I. Description of Blood Dyscrasias

Blood dyscrasias are a feature of over half of all drug-related deaths and include, but are not limited to, bone marrow aplasia, granulocytopenia, aplastic anemia, leukopenia, lymphoid hyperplasia, hemolytic anemia, and thrombocytopenia. All of these syndromes include pancytopenia to some degree.

*Bone marrow aplasia-* is defined as a profound loss of bone marrow resulting in pancytopenia. Drugs known to cause bone marrow aplasia include, but are not limited to, chloramphemicol, gold salts, mephenytoin, penicillamine, phenylbutazone, and trimethadione. In general these drugs are not first line therapy due to the rare occurance of marrow aplasia. Specific forms of aplasia include:

*Granulocytopenia-* is defined as a loss of polymorphonuclear neutrophils to a count lower than 500. Granulocytopenia primarily predisposes the patient to bacterial and fungal infections. Drugs known to cause granulocytopenia include, but are not limited to, captopril, cephalosporins, choral hydrate, chlorpropamide,

penicillins, phenothiazines, phenylbutazone, phenytoin, procainamide, propranolol, and tolbutamide.

*Aplastic anemia*- is a disorder involving an inability of the hematologic cells to regenerate and thus there is a dramatic depletion of one or more of the following cell types: neutrophils, platelets, or reticulocytes. Drugs associated with producing aplastic anemia are: 1) agents or compounds that produce bone marrow depression, for example cytotoxic drugs used in cancer chemotherapy; 2) agents or compounds that frequently, but inevitably, produce marrow aplasia, for example benzene; 3) agents or compounds that are associated with aplastic anemia, for example chloramphenicol, antiprotozoals, and sulfonamides.

Aplastic anemia is almost always a result of damage to the hematopoietic stem cells. There are two possible routes for the destruction of these cells: 1) direct damage to the stem cell DNA, and 2) cell cycle dependant depletion of later stage progenitor cells. In the first case, drugs or agents bind to and randomly damage the genetic material. This type of aplasia is associated with both early aplasia (immediate or direct cytotoxcity) or later myelodysplasia and leukemia. In the latter case, mitotically and metabolically active progenitor cells are preferentially affected and progenitor cell depletion may lead to unregulated proliferation of spared stem cells.

*Leukopenia*- is defined when the circulating peripheral white cell count falls below 5-10 x $10^9$ cells per liter. Circulating leukocytes consist of neutrophils, monocytes, basophils, eosinophils, and lymphocytes.

Neutropenia is defined when the peripheral neutrophil count falls below 2 X $10^9$ cells per liter. There are a number of drugs families that can cause neutropenia including, but not exclusive to, antiarrythmics (procainamide, propanolol, quinidine), antibiotics (chloramphenicol, penicillins, sulfonamides, trimethorpim-methoxazole, para-aminosalicyclic acid, rifampin, vancomycin, isoniazid, nitrofurantoin), antimalarials (dapsone, qunine, pyrimethamine), anticonvulsants ( phenytoin, mephenytoin, trimethadione, ethosuximide, carbamazepine), hypoglycemic agents (tolbutamide, chlorpropamide), antihistamines (cimetadine, brompheniramine, tripelennamine), antihypertensives (methydopa, captopril), antiinflammatory agents (aminopyrine, phenylbutazone, gold salts, ibuprofen, indomethacin), diuretics (acetazolamide, hydrochlorothiazide, chlorthalidone), phenothiazines (chlorpormazine, promazine, prochlorperazine), antimetabolite immunosuppresive agents, cytotoxic agents (alkylating agents, antimetabolites, anthracyclines, vinca alkyloids, cis-platinum, hydroxyurea, actinomycin D), and other agents (alpha and gamma interferon, allopurinol, ethanol, levamisole, penicillamine).

*Lymphoid hyperplasia*- is characterized by reactive changes within the T-cell regions of the lymph node that encroach on, and at times appear to efface, the germinal follicles. In these regions, the T-cells undergo progressive transformation to immunoblasts. These reactions are encountered particularly in response to drug-induced immunoreactivity. Drugs known to cause lymphoid hyperplasia are phenytoin, and mephenytoin.

*Hemolytic anemia*- is characterized by the premature destruction of red cells, accumulation of hemoglobin metabolic by-products, and a marked increase in erythroporesis within the bone marrow. Drugs know to cause hemolytic anemia include, but are not excluded to, methyldopa, penicillin, sulfonamides, and vitamin E deficiency.

*Thrombocytopenia*- is characterized by a marked reduction in the number of circulating platelets to a level below 100,000/mm$^3$. Drug-induced thrombocytopenia may result from decreased production of platelets or decreased platelet survival or both. Drugs known to cause thrombocytopenia include, but are not excluded to, ethanol, acetominophen, acetazolamide, acetylsalicyclic acid, 5-aminosalicylic acid, carbamazepine, chlorpheniramine, cimetadine, digitoxin, diltiazem, ethychlorynol, gold salts, heparin, hydantoins, isoniazid, levodopa, meprobamate, methyldopa, penicillamine, phenylbutazone, procainamide, quinidine, quinine, ranitidine, Rauwolfa alkaloids, rifampin, sulfonamides, sulfonylureas, cytotoxic drugs, and thiazide diuretics.

II. Impact of Stratification Based Upon Genotype in Drug Development for Drugs, Compounds, or Candidate Therapeutic Interventions that may Induce Blood Dyscrasias

Clozapine induced agranulocytosis is associated with differing HLA types and HSP70 variants in patients for whom responded to clozapine therapy but developed agranulocytosis. This is suggestive that a gene within the MHC region is associated with the manifestation of agranulocytosis in response to clozapine therapy. In a recent study, two ethnic groups were analyzed for genetic markers for the agranulocytosis. Tumor necrosis factor microsattelites d3 and b4 were found in higher frequencies in patients that experience clozapine-induced agranulocytosis. These data are suggestive that there is an involvement of tumor necrosis factor constellation polymorphism and clozapine-induced agranulocytosis.

There is evidence to suggest that there are safety response differences to drug therapy in reference to development of blood dyscrasias which may be attributable to genotypic differences between individuals. There is provided in this invention examples of gene pathways that are implicated in the disease process or its therapy and those that potentially cause this variability. The Detailed Description above

demonstrates how identification of a candidate gene or genes and gene pathways, stratification, clinical trial design, and implementation of genotyping for appropriate medical management of a given disease can be used to identify the genetic cause of variations in clinical response to therapy, new diagnostic tests, new therapeutic approaches for treating this disorder, and new pharmacuetical products or formulations for therapy. Gene pathways including, but not limited to, those that are outlined in the gene pathway Tables 1-6, preferably Table 3, and pathway matrix Table 8 and discussed below are candidates for the genetic analysis and product development using the methods described above.

Advantages of Inclusion of Pharmacogenetic Stratification in Clinical Development of Agents that May Cause Blood Dyscrasias

As an example of identification of the primary gene and relevant polymorphic variance that directly affects efficacy, safety, or both one could select an gene pathway as described in the Detailed Description, and determine the effect of genetic polymorphism and therapy efficacy, safety, or both within that given pathway. For example, referring to Table 8, genes involved in drug transport, phase I and phase II metabolism, protection from reactive intermediate damage, and immune responsiveness the optimization of therapy of by an agent known to have a blood dyscrasia side effect by determining whether the patient has a predisposing genotype in which the selected agents are more effective and or are more safe. In considering an optimization protocol, one could potentially predetermine the genotypic profile of these genes involved in the manifestation of the adverse effect, or those genes preeminently responsible for drug response. By embarking on the previously described gene pathway approach, it is technical feasibility to determine the relevant genes within such a targeted drug development program.

**Example 18**

Drug-Induced Toxicity: Cutaneous Toxicity

Drug–induced cutaneous toxicity includes, but is not excluded to, eczematous: photodermititis (phototoxic and photoallergic), exfoliative dermititis; maculopapular eruption; papulosquamous reactions: psoriaform, lichus planus, or pityriasis rosea-like; vesiculobullous reactions; txic epidermal necrolysis; pustular-acneform reactions; urticaria and erythemas: urticaria, erythema multiforme; nodular lesions: erythema nodosum, vasiculitis reaction; telangiectatic and LE reactions; pigementary reaction; other cutaneous reactions: fixed drug reactions, alopecia, hypertrichosis, macules, papules, angioedema, morbilliform-maculopapular rash, toxic epidermal necrolysis, erythema multiforme, erythema nodosum, contact

dermititis, vesicles, petechiae, exfolliative dermititis, fixed drug eruptions, and severe skin rash (Stevens-Johnson syndrome).

Drugs known to be associated with cutaneous toxicities include, but are not exclusive of, antineoplastic agents, sulfonamides, hydantoins and others listed for each type of toxicity.

*Uticaria and angioedema-* is defined as the transient apprearance of elevated, erythematous pruitic wheals (hives) or serpiginous exanthem. The apprearance of uticaria is percieved as ongoing immediate hypersensitivity reaction. Angioedema is defined as uticaria, but involving deeper dermal and subdermal sites. Uticaria and angioedema appear to result from dilation of local postcapillary venules. Degranulation of cutaneous mast cells may be involved.

Drugs associated with uticaria and angioedema include, but are not excluded to, antimicrobials include, but not exclusive of, 5-aminosalicylic acid, aminoglycosides, cephalosporins, ethambutol, isoniazid, metronidazole, miconazole, nalidixic acid, penicillins, quinine, rifampin, spectinomycin, sulfonamides, and other drugs: asparaginase, aspirin and other non-steroidal antiinflammatory agenets, calcitonin, chloral hydrate, chlorambucil, cimetidine, cyclophosphamide, daunorubicin, ergotamine, ethchlorvynol, doxorubicin, ethosuximide, ethylenediamine, glucocorticoids, melphalan, penicillamine, phenothiazines, procainamide, procarbazine, quinidine, tartazine, thiazide diuretics, thiotepa.

*Morbilliform-maculopapular rash-* are rashes that result in eruptions or are morbilliform in nature.

Drugs associated with rashes include, but are not limited to, 5-aminosalicyclic acid, cephalosporins, erythromycin, gentamicin, penicillins, streptomycin, sulfonamides, allopurinol, barbiturates, captopril, coumarin, gold salts, hydantoins, thiazide diuretics.

*Toxic epidermal necrolysis and erythroderma and exfoliative dermititis-*

Cutaneous erythroderma, edema, scaling, and fissuring may occur in response to certain drugs. Drugs associated with these types of cutaneous reactions include, but are limited to, allopurinol, amikacin, captopril, carbamazepine, chloral hydrate, chlorambucil, chloroquine, chlorpromazine, cyclosporine, diltiazem, ethambutol, ethylenediamine, glutethimide, gold salts, griseofulvin, hydantoins, hydroxychloroquine, minoxidil, nifedipine, nonsteroid antiinflammatory agents, penicillin, phenobarbital, rifampin, spironolactone, sulfonamides, trimethadione, trimethoprim, tocainamide, tocainide, vancomycin, verpamil.

*Erythema mutliforme-* is characterized by a hypersensitivity reaction in blood vessels of the dermis. The hypersensitivity is the result of immune complexes formed by small molecules interacting with proteinaceous components of the blood

vessels. In cases whereby the mucosal membranes of the mouth and eye are involved, is referred to as Stevens-Johnson syndrome. Typically the cutaneous lesions, blisters and painful erosions occur in the mout and eye.

Drugs associated with erythema mulitforme include, but are not limited to, allopurinol, acetominophen, amikacin, barbiturates, carbamazepine, chloroquine, chlorporamide, clindamycin, ethambutol, ethosuximide, gold salts, glucocorticoids, hydantoins, hydralazine, hydroxyurea, mechlorethamine, meclofenamate, penicillins, phenothiazides, phenophthalein, phenylbutazone, rifampin, streptomycin, sulfonamides, sulfonylureas, sulindac, vaccines.

*Fixed drug eruptions-*

Drug associated with fixed drug eruptions include, but are not excluded to, acetominophen, 5-aminosalicyclic acid, aspirin, barbiturates, benzodiazepines, barbiturates, chloroquine, dapsone, dimethylhydrinate, gold salts, hydralazine, hyoscine, ibuprofen, iodides, meprobamate, methanamine, metronidazole, penicillins, phenobarbital, phenolphthalein, phenothiazides, phenylbutazone, procarbazine, pseudoephedrine, quinine, saccharin, streptomycin, sulfonamides, and tetracyclines.

*Erythema nodosum-* is an innflammatory reaction in subcutaneous fat which represents a hypersentivity reaction to a number of antigenic stimuli. Multiple red, painful nodules do not ulcerate but involute and leave a yeloow-purple bruises. Small molecules intreracting with proteinaceous components forma asenstitizing antigen.

Drugs associated with producing erythema nododum include, but are not excluded to, bromides, oral contraceptives, penicillins, and sulfonamides.

*Contact dermititis-* is characterized by eruptions on histological analysis to epidermal intercellular edema (spongiosis). Contact dermititis can be caused by allergic or irritant mechanisms. Allergic contact dermititis is a delayed hypersensitivity reaction that can occur in response to a variety of small molecules that when bound to proteinaceous components of the skin form a sensitizing antigen. The antigen is processed by Langerhans' cells in the epidermis, presenting the antigen to the circulating T lymphocytes. Irritant dermititis is produced by substances that irritate or have a direct toxic effect on the skin.

Drugs associated with contact dermititis side effects include, but are not limited to, ambroxol, amikacin, antihistamines, bacitracin, benzalkonium chloride, benzocaine, benzyl chloride, cetl alcohol, chloramphenicol, chlorpormazine, clioquinol, colophony, ethylenediamine, fluorouracil, formaldehyde, gentamycin, glucocorticoids, glutaraldehyde, heparin, hexachlorophene, iodochlorhydroxyquin, lanolin, local anesthestics, minoxidil, naftin, neimycin, nitrofurazone, opiates, para-

aminobenzoic acid, parabens, penicillins, phenothiazines, prolflavine, propylene glycol, streptomycin, sulfonamides, thimerosal, timolol.

*Impact of Stratification Based Upon Genotype in Drug Development for Drugs, Compounds, or Candidate Therapeutic Interventions that May Induce Cutaneous Reactions*

Recently, it has been decribed that there is a deletion polymorphism in the B2 bradykinin receptor gene (B2BKR). It was revealed that there is a 9 base pair deletion in exon 1 of the B2BKR gene and upon inspection of patients experienceing angioedema, patients with immunochemical evidence of angioedema were homozygous for no deletion at that site. These results were suggestive of B2BKR genotype influence on the clinical status and manifestation angioedema.

There is evidence to suggest that there are safety response differences to drug therapy in reference to development of cutaneous reactions which may be attributable to genotypic differences between individuals. There is provided in this invention examples of gene pathways that are implicated in the disease process or its therapy and those that potentially cause this variability. The Detailed Description above demonstrates how identification of a candidate gene or genes and gene pathways, stratification, clinical trial design, and implementation of genotyping for appropriate medical management of a given disease can be used to identify the genetic cause of variations in clinical response to therapy, new diagnostic tests, new therapeutic approaches for treating this disorder, and new pharmacuetical products or formulations for therapy. Gene pathways including, but not limited to, those that are outlined in the gene pathway Tables 1-6, more preferably Table 3, and pathway matrix Table 8 and discussed below are candidates for the genetic analysis and product development using the methods described above.

Advantages of Inclusion of Pharmacogenetic Stratification in Clinical Development of Agents that May Cause Cutaneous Reactions

As an example of identification of the primary gene and relevant polymorphic variance that directly affects efficacy, safety, or both one could select an gene pathway as described in the Detailed Description, and determine the effect of genetic polymorphism and therapy efficacy, safety, or both within that given pathway. For example, referring to Table 8, genes involved in drug transport, phase I and phase II metabolism, protection from reactive intermediate damage, and immune responsiveness, the optimization of therapy of by an agent known to have a cutaneous side effect by determining whether the patient has a predisposing

genotype in which the selected agents are more effective and or are more safe. In considering an optimization protocol, one could potentially predetermine the genotypic profile of these genes involved in the manifestation of the adverse effect, or those genes preeminently responsible for drug response. By embarking on the previously described gene pathway approach, it is technical feasibility to determine the relevant genes within such a targeted drug development program.

**Example 19**

Drug-Induced CNS Toxicity

Drug-induced central nervous system toxicity includes CNS stimulation or CNS depression. Characteristics of CNS toxicity include, but are not limited to, tinnitus and dizziness, acute dystonic reactions, parkinsonian syndrome, coma, convulsions, depression and psychosis, sweating, mydriasis, hyperpyrexia, centrally mediated cardiovascular involvement (hypertension, tachycardia, extrasystoles, arrythmias, circulatory collapse) and respiratory depression or tachypnea. Drugs known to be associated with CNS toxicity include, but are not exclusive of, salicylates, antipsychotics, sedatives, cholinergics,

*Impact of Stratification Based Upon Genotype in Drug Development for Drugs, Compounds, or Candidate Therapeutic Interventions that May Induce CNS Toxicity*

There is evidence to suggest that there are safety response differences to drug therapy in reference to development of CNS toxicities which may be attributable to genotypic differences between individuals. There is provided in this invention examples of gene pathways that are implicated in the disease process or its therapy and those that potentially cause this variability. The Detailed Description above demonstrates how identification of a candidate gene or genes and gene pathways, stratification, clinical trial design, and implementation of genotyping for appropriate medical management of a given disease can be used to identify the genetic cause of variations in clinical response to therapy, new diagnostic tests, new therapeutic approaches for treating this undesirable adverse effect, and new pharmacuetical products or formulations for therapy. Gene pathways including, but not limited to, those that are outlined in the gene pathway Tables 1-6, more preferably Table 3, and pathway matrix Table 8 and discussed below are candidates for the genetic analysis and product development using the methods described above.

Advantages of Inclusion of Pharmacogenetic Stratification in Clinical Development of Agents that May Cause CNS Toxicities

As an example of identification of the primary gene and relevant polymorphic variance that directly affects efficacy, safety, or both one could select an gene pathway as described in the Detailed Description, and determine the effect of genetic polymorphism and therapy efficacy, safety, or both within that given pathway. For example, referring to Table 8, genes involved in drug transport, phase I and phase II metabolism, protection from reactive intermediate damage, the optimization of therapy of by an agent known to inpart CNS toxic or undesirable side effect or effects by determining whether the patient has a predisposing genotype in which the selected agents are more effective and or are more safe. In considering an optimization protocol, one could potentially predetermine the genotypic profile of these genes involved in the manifestation of the adverse effect, or those genes preeminently responsible for drug response. By embarking on the previously described gene pathway approach, it is technical feasibility to determine the relevant genes within such a targeted drug development program.

**Example 20**

Drug-Induced Liver Toxicity

Drug-induced liver disease or drug-induced liver toxicity can manifest as zonal necrosis, nonspecific focal hepatitis, viral hepatitis-like reactions, inflammatory or noninflammatory cholestasis, small or large droplet fatty liver, granulomas, chronic hepatitis, fibrosis, tumors, or vascular lesions.

In the majority of the cases of known drug-induced liver toxicity, the drug is metabolized to a form that is deleterious to hepatic, or extrahepatic function. There are many endogenous or exogenous compounds that may be considered to attenuate or ablate toxic hepatocyte-produced metabolite mechanisms or effects of hepatic or extrahepatic damage.

In hepatocellular damage, free oxygen radicals may be generated in the hepatic metabolic processes that are deleterious to intracellular organelles, DNA, or metabolic pathways. There are endogenous cytoprotective agents that may prevent free radical-mediated damage such as retinoids, flavins, reduced glutathione, vitamin E, S-adenylylmethionine, and the enzyme superoxide dismutase (SOD). In animal models in which SOD activity is diminished or absent, the liver function was normal, but the sensitivity to toxin challenge was heightened.

In cholestatic damage, the bile salt uptake, metabolism, secretion, or transport is compromised and the residual increased bile salt concentrations are deleterious to hepatocyte function. The increase in bile salts is the main metabolic

disturbance that initially leads to jaundice and pruritis and can progress to pancreatitis, hyperbilirubinemia, biliary cirrhosis, and hepatic encephalopathy.

In both cases of drug-induced liver toxicity, the drug must first be absorbed and enter in the hepatic circulation. Further, clinically it is often difficult to determine whether cholestatic damage leads to hepatocellular damage or whether hepatocellular damage leads to cholestatic damage. In many cases, until the patient is symptomatic, the underlying damage mechanisms may be clinically overlooked. By the time the drug-induced liver disease is symptomatic, the damage, be it hepatocellular or cholestatic or both, may be irreversible.

*Identification of Genes involved in Drug-Induced Liver Toxicity*

Thus, in the process of identifying drug- or xenobiotic-induced liver toxicity, one skilled in the art would identify key metabolic enzymes or bile cannicula transport processes that would be linked with either hepatocellular damage or cholestasis or combination of hepatocellular damage or cholestasis.

Hepatocellular damage may be the result of direct chemical mediated effects, may be severe, and usually is associated with damage within organelles, DNA and membranes. Clinically there is a marked elevation of SGOT and SGPT as well as other enzymes. In cases of cholestasis there is jaundice, pruritis, a marked elevation of bile salts and alkaline phosphatase activity, but not an elevation of SGOT or SGPT. In cases of toxic liver disease there is difficulty, at least initially to determine the underlying etiology. Clinically, symptoms may not appear as clear as described above. Further, depending on the rate and extent of the damage, hepatocellular damage may be masked or asymptomatic until liver impairment has induced cholestasis.

Potentially hepatotoxic agents can be divided broadly into two groups: intrinsic hepatotoxins and idiosyncratic hepatotoxins. Intrinsic hepatotoxins produce acute liver damage in a predictable, dose-dependent fashion shortly after ingestion or exposure. Generally, all subjects exposed will uniformly exhibit signs and symptoms. In this category, the effects seen in humans can be mimicked in animal models. Examples of intrinsic hepatotoxins are carbon tetrachloride, 2-nitropropane, trichloroethane, the octapeptide toxins of the Amanita mushroom species, and the antipyretic, acetominophen. In some of these cases, toxic metabolites result in covalent modification of hepatocyte macromolecules or reactive oxygen intermediates leads to peroxidation of cell membrane lipids or other intracellular molecules.

In contrast, idiosyncratic hepatotoxins produce liver damage in an unpredictable, dose-independent manner after a latent period of ingestion or

exposure. Animal models or experimental data is generally incapable of predicting the effect in humans. Further, idiosyncratic hepatotoxins do not uniformly affect a population; a subset of the group exposed may or may not exhibit signs or symptoms. Range of symptoms are from mild to severe and is thought to coincide with differences in the pathways of drug or xenobiotic biotransformation or immune-mediated drug sensitivity (drug allergy). In idiosyncratic drug-induced liver disease, fever, arthralgias, rash, eosinophilia, are often prominent and indicate a hypersensitivity reaction.

*Impact of Stratification Based Upon Genotype in Drug Development for Drugs, Compounds, or Candidate Therapeutic Interventions that may Induce Hepatotoxicity*

Genes encoding proteins with catalytic function that are involved in the metabolism of drugs or xenobiotics are listed in Tables 3 and 8 below. Further listed are those proteins that are involved in the uptake, transport, or secretion into the bile cannicula. Below are further specific example of drug-specific effects on the liver.

*Acetaminophen-Induced Liver Disease*

Acetominophen is a readily available, easy to administer analgesic that is an example of a intrinsic hepatotoxin. This hepatotoxin causes zonal necrosis and acute liver failure and is associated with renal failure. Although a high dose (10-15 grams) is required for significant liver injury to occur, the onset of initial symptoms does not occur until hours after ingestion. The progression of symptoms occurs including progressive liver failure with hepatic encephalopathy, prolongation of prothrombin time, hypoglycemia, and lactic acidosis. The liver injury is caused by a toxic metabolite of acetominophen via the P450 metabolizing system. This toxic intermediate at low concentrations is conjugated with glutathione. However, in toxic doses, the conjugating enzymes stores are exhausted and the reactive intermediate reacts with intracellular proteins and results in cellular dysfunction and ultimately death. The rate of metabolism is dependent on the concentrations of both P450 and glutathione. Speeding this toxic pathway may include increasing the available P450 or reducing the availablility of glutathione, e.g. using known inducers of P450 such as ethanol and and phenobarbital; and known inhibitors of glutathione concentrations, e.g., ethanol and fasting. Acetominophen toxicity is completely reversed if the drug is removed. Chronic ingestion may produce subclinical liver injury, centrilobular necrosis, or chronic hepatitis; however all reversible if the drug is removed.

*Amiodarone-Induced Liver Disease*

Amiodarone is used in treatment of refractory arrythmias. In some patients amiodarone produces mild to moderate increases of serum transaminases which are generally accompanied by engorgement of lysosomes with phospholipid. In a fraction of the patients, a more severe liver injury develops which histologically resembles alcoholic hepatitis: fat infiltration of hepatocytes, focal necrosis, fibrosis, polymorphonuclear leukocyte infiltrates, and Mallory bodies. The lesion may progress to micronodular cirrhosis, with portal hypertension and liver failure. Hepatomegaly is seen, but jaundice is rare.

Amiodarone accumulates in lysosomes and inhibits lysosomal phopholipases, however the connection between this mechanism and alcoholic hepatitis histopathology is unknown. Unfortunately, rapid discontinuation of amiodarone increases the risk of cardiac arrythmias.

*Chlopromazine-Induced Liver Disease*

Chlorpromazine is an anti-psychotic agent which, in a small portion of the patient population can produce a cholestatic reaction. Symptoms include fever, anorexia, arthalgias, pruritis, jaundice, and eosinophilia is common. This idiosyncratic type of liver toxicity suggests a hypersensitivity type reaction. The symptoms subside over a period of weeks following discontinuation, Rarely, residual cholestatic disease occurs, treatment for pruritis and fat-soluble vitamin supplementation may be required, but eventual recovery almost always occurs.

*Erythromycin-Induced Liver Disease*

Erythromycin, a broad spectrum antibiotic, can be accompanied by a cholestatic reaction. Inflammatory cell infiltration and liver cell necrosis may occur. The hepatoptoxicity presents as right upper quadrant pain, fever, and variable cholestatic symptoms. The prognosis is uniform and will occur after readminstration of the drug, The mechanism of action is unknown.

*Halothane-Induced Liver Disease*

Halothane is a gaseous anthesthetic and can, in rare instances, cause a viral-like hepatitis syndrome. In severe cases, this hepatotoxicity, may cause fatal massive heaptic necrosis. Severe reactions seem to appear after previous or multiple exposure to halothane. It is known that the P450 metabolites of this xenobiotic are responsible for the mechanism of hepatic injury.

*Isoniazid (INH)-Induced Liver Disease*

Isoniazid is used as a single drug in the prophylaxis of tuberculosis. In 10-20% of of the persons taking INH, subclinical liver injury occurs. The conversion of INH to acetylhydrazine is via acetylation. In slow acetylators, INH is more hepatotoxic. The conversion of INH to acetylhydrazine to diacetylhydrazine is impaired. In slow acetylators, the acetylhydrazine is not well metabolized and is further oxidized by one of the P450 enzymes to a toxic, reactive molecule that is responsible for the liver disease. Discontinuation of the drug returns the enzymatic levels to normal and the liver is able to restore activity.

*Sodium Valproate-Induced Liver Disease*

Sodium valproate is an anti-epileptic agent that is routinely prescribed for petit mal epilepsy and in some cases produces severe hepatotoxicity. Similar to INH, sodium valproate is accompanied by a high incidence of transient, slight and asymptomatic increases in serum transaminases. Usually the increased enzyme activity appears after weeks of treatment. In rare cases of severe liver toxicity, the nonspecific systemic and digestive symptoms are followed by jaundice, evidence of liver failure, as well as encephalopathy and coagulopathy. The mechanism of hepatotoxicity is unknown, however there are theories that there is impairment of mitochondiral oxidation of long-chain fatty acids by a metabolite of the parent drug. Symptoms subside with little to no residual liver dysfunction after discontinuing the drug.

*Oral Contraceptive Induced Liver Disease*

Estrogen, progesterone, and combination oral contraceptives can produce several adverse effects on the heptobiliary system. They are 1) hepatocellular cholestasis, 2) liver cell neoplasias, 3) increased predisposition to cholesterol and gall stone formation, 4) hepatic vein thrombosis. These cholestatic hepatotoxic effects are attributed to estrogen's direct effect on bile formation. The mechanism of action is unknown.

There is evidence to suggest that there are safety response differences to drug therapy in reference to development of drug-induced liver toxicity which may be attributable to genotypic differences between individuals. There is provided in this invention examples of gene pathways that are implicated in the disease process or its therapy and those that potentially cause this variability. The Detailed Description above demonstrates how identification of a candidate gene or genes and gene pathways, stratification, clinical trial design, and implementation of genotyping for appropriate medical management of a given disease can be used to identify the genetic cause of variations in clinical response to therapy, new diagnostic tests, new

therapeutic approaches for treating this disorder, and new pharmacuetical products or formulations for therapy. Gene pathways including, but not limited to, those that are outlined in the gene pathway Tables 1-6, more preferably Table 3, and pathway matrix Table 8 and discussed below are candidates for the genetic analysis and product development using the methods described above.

Advantages of Inclusion of Pharmacogenetic Stratification in Clinical Development of Agents that May Cause Liver Toxicity

As an example of identification of the primary gene and relevant polymorphic variance that directly affects efficacy, safety, or both one could select an gene pathway as described in the Detailed Description, and determine the effect of genetic polymorphism and therapy efficacy, safety, or both within that given pathway. For example, referring to Table 8, genes involved in drug transport, phase I and phase II metabolism, excretion, hepatic cannicular uptake and concentration, and protection from reactive intermediate damage the optimization of therapy by an agent known to have a hepatic side effect by determining whether the patient has a predisposing genotype in which the selected agents are more effective and or are more safe. In considering an optimization protocol, one could potentially predetermine the genotypic profile of these genes involved in the manifestation of the adverse effect, or those genes preeminently responsible for drug response. By embarking on the previously described gene pathway approach, it is technical feasibility to determine the relevant genes within such a targeted drug development program.

**Example 21**

Drug-Induced Cardiovascular Toxicity

Drug induced cardiovascular toxicities include but are not excluded to arrhythmias, tachycardia, extrasystoles, circulatory collapse, QT prolongation, cardiomyopathy, hypotension, or hypertension. Drugs known to elicit these type of responses include but are not excluded to theophylline, hydantoins, doxorubicin, daunorubicin.

Arrythmias-If the normal sequence of electrical impulse and propagation through myocardial tissue is perturbed, an arrythmia occurs. Broadly, arrythmias fall into one of three categories: bradyarrythmias (slowing or failure of the initiating impulse), heart block (an impaired propagation through node tissue or atrial or ventricular muscle), and tachyarrythmias (abnormal rapid heart rhythms). Subcategories include: sinus bradycardia, atrioventricular block (AV block), sinus tachycardia, ventricular tachycardia, atrial flutter, multifocal atrial tachycardia, polymorphic ventricular tachycardia with or without QT prolongation, frequent or

difficult to terminate ventricular tachycardia, atrial tachycardia with or without AV block, ventricular bigeminy, and ventricular fibrillation. Drugs known to induce these types of arrythmias include, but are not excluded to, digitalis, verapamil, diltiazem, b-adrenergic blockers, clonidine, methyldopa, quinidine, flecainide, propafenone, theophylline, sotalol, procainamide, disopyramide, certain non-cardioactive drugs ( ), and amiodarone.

Heart Rate, Tachycardia-Heart rate is under both sympathetic and parasympathic control. The influence of heart rate on cardiac output is paramount. Drugs affecting heart rate include, but are not limited to, sympathomimetics, parasympathomimetics, and agents or compounds affecting these two central inputs.

Extasystoles- is defined as premature myocardial excitation. Extrasystoles can include atrial, nodal, or ventricular. Other asynchronous pathologies may result from these systoles. Drugs known to be associated with extra systoles include, but are not excluded to, agents that prolong the depolarization time, agents that leave a residual available intracellular calcium, or agents that alter the function of the K+ or Na+ channel activity.

QT Prolongation- is the interval on an electrocardiogram that indicates ventricular action potential duration. QT prolongation can lead to uncoordinated atrial and ventricular action potentials. In these circumstances of delayed or prolonged polymorphic ventricular afterdepolarizations, resultant abnormal triggering of secondary, uncoordinated depolarizations can occur. Two of these conditions are explained as follows and may be associated with underlying rapid or slow heart rate: 1) under conditions of residual excess intracellular calcium (myocardial ischemia, adrenergic stress, digitalis intoxication), and 2) under conditions of marked prolongation of cardiac action potential (agents (antiarrythmics or others) that prolong action potential duration).

Cardiomyopathy-There are broadly three categories of cardiomyopathies: dilated, hypertrophic, and restrictive. These cardiac muscular diseases can be of mechanical or acquired origin.

Dilated cardiomyopathies are generally caused by myocardial injury that results in depressed systolic function and progressive ventricular dilatation. Drug induced dilated cardiomyopathy can occur in the presence of, but are not excluded to, ethanol, chenotherapeutic agents, elemental compounds, and catecholamimetics.

Hypertrophic cardiomyopathy is the presentation of grossly assymetric (eccentric) or symmetric (concentric) hyoertrophy of the left ventricle in the absence of another cardiac or systemic disease capable of producing the disproportionate increase in ventricle mass. In drug induced hypertrophic cardiomyopathy, there may be compensatory hypertrophy of the left ventricle in response to inordinate and or

sustained hypertension or prolonged reduced or insufficient cardiac output as a result of myocardial injury or noncardiac mediated physiological events.

Restrictive cardiomyopathies are the result of a primary abnormality of diastolic function (impaired filling). Impaired diastolic function can occur as a result of morphologically detectable myocardial or endomyocardial disease, interstitial deposition of deposition of abnormal substances (infiltrative), intracellular accumulation of abnormal substances (strage diseases), or as a result of endomyocardial disease. In the last category, anthracyclines have been associated with both dilated and restrictive cardiomyopathies.

Blood Pressure-Blood pressure is regulated in a complex interplay of neural and endocrine mechanisms. These mechanisms are aimed at the physiologic contorl of cardiac output, delivery of blood components to the tissues, and removal of metabolic by-products from the tissues.

Hypertension is defined as the elevated arterial blood pressure either an increase of systolic or diastolic pressure or both. Secondary hypertension can be associated with drugs and chemicals including, but not limited to, cyclosporine, oral contraceptives, glucocorticoids, mineralocorticoids, sympathomimetics, tyramine, and MAO inhibitors.

Hypotension is defined as the reduction in blood pressure that is associated with orthostatic hypotension, syncope, head injury, hepatic failure, antidiuresis, myocardial infarction and cardiogenic shock. Drug-induced hypotension is associated drugs including, but not exclusive of, parasympathomimetics, diuretics, and direct acting cardiac agents.

Advantages of Inclusion of Pharmacogenetic Stratification in Clinical Development of Agents that May Cause Cardiovascular Toxicity

As an example of identification of the primary gene and relevant polymorphic variance that directly affects efficacy, safety, or both one could select an gene pathway as described in the Detailed Description, and determine the effect of genetic polymorphism and therapy efficacy, safety, or both within that given pathway. For example, referring to Table 8, genes involved in drug transport, phase I and phase II metabolism, and protection from reactive intermediate damage the optimization of therapy of by an agent known to have a cardiovascular side effect by determining whether the patient has a predisposing genotype in which the selected agents are more effective and or are more safe. In considering an optimization protocol, one could potentially predetermine the genotypic profile of these genes involved in the manifestation of the adverse effect, or those genes preeminently

responsible for drug response. By embarking on the previously described gene pathway approach, it is technical feasibility to determine the relevant genes within such a targeted drug development program.

**Example 22**

Drug-Induced Pulmonary Toxicity

Drug induced pulmonary toxicity includes, but is not excluded to, asthma, acute pneumonitis, eosinophilic pneumonitis, fibrotic and pleural reactions, and interstitial fibrosis. Drug know to elicit pulmonary toxicity include, but are not excluded to, salicylates, nitrofuratoin, busulfan, nitrofurantoin, and bleomycin.

Advantages of Inclusion of Pharmacogenetic Stratification in Clinical Development of Agents that May Cause Pulmonary Toxicities

As an example of identification of the primary gene and relevant polymorphic variance that directly affects efficacy, safety, or both one could select an gene pathway as described in the Detailed Description, and determine the effect of genetic polymorphism and therapy efficacy, safety, or both within that given pathway. For example, referring to Table 8, genes involved in drug transport, phase I and phase II metabolism, excretion, protection from reactive intermediate damage, and immune responsiveness, the optimization of therapy of by an agent known to have a pulmonary side effect by determining whether the patient has a predisposing genotype in which the selected agents are more effective and or are more safe. In considering an optimization protocol, one could potentially predetermine the genotypic profile of these genes involved in the manifestation of the adverse effect, or those genes preeminently responsible for drug response. By embarking on the previously described gene pathway approach, it is technical feasibility to determine the relevant genes within such a targeted drug development program.

**Example 24**

Drug-Induced Renal Toxicity

Drug-induced renal toxicity includes, but is not exclusived to, glomerulonephritis and tubular necrosis. Drugs associated with eliciting renal toxicity include, but are not excluded to, penicillamine, aminoglycoside antibiotics, cyclosporine, amphotericin B, phenacetin, and salicylates.

Advantages of Inclusion of Pharmacogenetic Stratification in Clinical Development of Agents that May Cause or are Associated with Renal Toxicity

As an example of identification of the primary gene and relevant polymorphic variance that directly affects efficacy, safety, or both one could select an gene pathway as described in the Detailed Description, and determine the effect of genetic polymorphism and therapy efficacy, safety, or both within that given pathway. For example, referring to Table 8, genes involved in drug transport, phase I and phase II metabolism, and renal tubular uptake and concentration the optimization of therapy of by an agent known to have a renal side effect by determining whether the patient has a predisposing genotype in which the selected agents are more effective and or are more safe. In considering an optimization protocol, one could potentially predetermine the genotypic profile of these genes involved in the manifestation of the adverse effect, or those genes preeminently responsible for drug response. By embarking on the previously described gene pathway approach, it is technical feasibility to determine the relevant genes within such a targeted drug development program.

**Example 24**

Asthma

*I. Description of Asthma*

Asthma can be an acute or chronic condition associated with inflammation of the lower airways and variable levels of airflow obstruction. Asthma symptoms vary among individuals and may include wheezing, shortness of breadth, tightness of the chest, trouble controlling a cough, persistent cough at night, difficulty breathing during or soon after physical exertion or exercise, or waking up at night due to one of these symptoms. Episodes of these symptoms (referred to as asthma attacks, flare-ups, or exacerbations) occur when there is sufficiently severe airway constriction to render a patient almost unable or unable to breathe. There can be warning signs , however, many attacks are sudden and unanticipated.

Individuals with asthma have inflamed airways that are supersensitive to inducers of asthma which exacerbate asthma and enhance underlying inflammation such as allergens, respiratory infections, or industrial pollutants. Provokers of asthma leading to bronchospasm include exercise or physical activities, irritants, emotions and aspirin. Asthma attacks are associated with swollen and inflamed linings of the airways, excess mucus in the airways, and bronchospasm which are reversible. In chronic asthma, there is persistent activation of resident cells (e.g. basophils, eosinophils, neutrophils) lining the airway leading to chronic inflammation which can result in irreversible changes in the airway pasages. These permanent changes are part of a remodeling process.

Recent evidence has suggested that airway inflammation is a major factor in the pathogenesis and in the severity of the disease. One theory holds that asthsma is a T helper 2 (Th2) cell-driven chronic eosinophilia mediated via dendritic and other antigen-presenting cells. The inflammatory nature of the disease is multicellular in nature, with mast cells, eosinophils, macrophages, basophils, lymphocytes, neutrophils, and epithelial cells participating and therefore immunoglobins, cytokines, chemokines, adhesion molecules, proteinases, inflammatory mediators, and growth factors are involved in various stages and interact to maintain and amplify the inflammatory response. The net result of these interactions is persistent inflammation and repair, ultimately leading to irreversible airway remodeling.

*II. Current therapies for Asthma*

Because asthma results from a complex combination of mediators of inflammation, most useful anti-asthma agents affect pathways for these mediators. In acute or chronic asthma, the therapeutic categories include: immunosuppressive agents including glucocorticoids, antiinflammatory agents including leukotriene receptor agonists and mast cell stabilizers (cromolyn sulfate); bronchodilators including β-adrenergic agonists, sympathomimetic agents, and xanthines; and agents to treat cough and excess mucus including expectorants and mucolytics.

Corticosteroids affect the inflammation within the airways by decreasing growth and development of mast cells, inducing apoptosis, suppressing lymphocyte generation of IL-5 and other cytokines, inhibiting some mediator release, inhibiting cytokine production, inhibiting the transcription of cytokines (for example IL-8, TNF-α, prototypic antiviral chemokine (regulated-on-activation normal T-expressed and secreted, RANTES), and GM-CSF), and inhibiting nitric oxide synthesis.

β-Adrenergic agonists and sympathomimetics affect the pulmonary airway lining in a well-characterized mechanism of β-adrenergic receptor activation of adenylyl cyclase as well as cAMP independent mechanisms. Bronchodilation is the immediate clinical effect.

Leukotriene modifiers affect the airway by inhibition of 5-lipoxygenase, the initial enzyme of leukotriene biosynthesis, and exert their effect by decreasing leukotriene production, thereby interferring with eosinophil migration and other processes.

Corticosteroids affect the inflammation within the airways by decreasing growth and development of mast cells, inducing apoptosis, suppressing lymphocyte generation of IL-5 and other cytokines, inhibiting some mediator release, inhibiting cytokine production, inhibiting the transcription of cytokines (for example IL-8,

TNF-α, prototypic antiviral chemokine (regulated-on-activation normal T-expressed and secreted, RANTES), and GM-CSF), and inhibiting nitric oxide synthesis.

Corticosteroids in combination with long-acting β-adrenegic agonists work well as combination therapy.

Cromones are believed to act on the airway by modifying mediator release, and inhibiting mast cell degranulation.

Xanthines are believed to act on the airway in asthma by inhibiting eosinophil cell migration, and enhancing β-adrenergic pathway mediated bronchodilation via the inhibition of phosphodiesterase.

Difficult to treat or therapy-resistant asthma syndromes present a challenge to clinicians. They include difficult acute and chronic, as well as chronic severe, acute severe, therapy-resistant, difficult to control and corticosteroid-dependent asthma.

*III. Limitations of Current Therapies for Asthma*

Limitations Involving Efficacy

The therapies discussed above do not reverse the underlying pathological process in asthma; they merely slow or retard the progression of asthma. As thickening of the airways occurs and becomes irreversible the therapeutic options become limited. Thus, therapies for asthma are aimed at reduction of inflammatory processes and control of symptoms starting at the earliest date (frequently in the pediatric setting).

The limitations of the adrenegic agonist compounds used for the treatment of asthma include short duration of action and ligand desensitization. Excessive use of short acting β-adrenergic agonists has been proposed to lead to loss of asthma control and consequent increases in morbidity and mortality. Long acting bronchoactive/bronchoprotective agonists acting at adrenergic receptors have supplanted short duration β-agonists.

Short-acting β-adrenergic agonists are primarily used for the relief of acute asthma symptoms. Excessive reliance on these agents is generally not advisable because 1) β-adrenergic receptors undergo a rapid desenitization and the agonist becomes an ineffective bronchodilator, and 2) repetitive high doses of short acting β-adrenergic agonists may be detrimental to to the control of asthma by potentially interfering with corticosteroid action. This desensitization occurs through a process involving G-protein receptor coupled-kinases and or cAMP dependent protein kinase or by enhanced degradation of cAMP by phosphodiesterase activity.

Glucocorticoid associated side effects include increased appetite, weight gain, fluid retention, acne, ecchymosis, development of Cushingoid facies,

hypertension, hyperkalemia, diabetes, hyperglycemia, hyperosmolar state, hyperlipidemia, hepatic steatosis, atherosclerosis, myopathy, aseptic necrosis, osteoporosis, ulcers, pancreatitis, psuedotumor cerebri, pyschosis, glaucoma, cataract formation, vascular necrosis, increased suseptibility to infection, impairment of the hypothalamus-pituitary axis, decreased thyroid hormone serum binding protiens, and impaired wound healing.

Theophylline or other phosphodiesterase inhibitors have been shown to have a narrow therapeutic window and can result in life-threatening cardiac arrythmias.

Difficult to treat asthma involves a spectrum of disease that responds suboptimally to doses of glucocorticoids. In the face of partial response to inhaled or oral steroids, higher doses are administered risking steroid associated side-effects.

The reduction of clinical symptoms of asthma following antiinflammatory therapy may only become evident after several weeks to months of therapy. The slow action of these therapies creates problems for the clinician seeking to expeditiously determine optimal therapy for an individual patient. The development of genetic tests to predict response to different agets will allow selection of optimal therapy with less of the time consuming empirical clinical decision making required presently.

Limitations Involving Toxicity or Undesired Side Effects

There are toxicities and undesired side effects associated with the above current therapies for asthma that require monitoring. Drugs used to treat asthma may cause death, disability, disease, and place a fetus at risk. The undesired side effects or toxicities are listed for each drug category as described above.

*IV. Impact of Stratification Based Upon Genotype in Drug Development for Drugs, Compounds, or Candidate Therapeutic Interventions for Asthma*

In a recent report, it was demonstrated that the 5-lipoxygenase (5-LO) gene promoter variation among asthma patients is linked to drug response to 5-LO inhibitors (Drazen et al., Nature Genetics 22: 1999). In a clinical trial to test efficacy of a potent, selective 5-LO inhibitor (ABT-761), the trial was abruptly closed due to inordinate event rate of abnormal liver function tests. Although the projected enrollment was not reached, the interim data suggested superior efficacy regarding forced expiratory volume in the high dose relative to low dose or placebo groups. The investigators chose to stratify the high dose and placebo group of the enrolled patients based upon genotype of the 5-LO gene promoter. The 5-LO gene promoter has been found to contain 3-6 tandem repeats of the Sp1-binding motif. The wild-type allele was designated as 5 tandem repeats and had a frequency of

0.772 in the study population. The forced expiratory volume data indicated that heterozygous patients on high-dose active treatment had, on average, an improvement of forced expiratory volume within one week (23.3 ± 6.0%) and was similar to the wild-type patients (18.8 ± 3.6%). In contrast, the patients with mutant genotype had no benefit from active 5-LO inhibitor treatment (-1.2 ± 2.9%). In the table below, the trial outcome data is described for two periods following treatment with high dose or placebo.

| | $FEV_1$, % change from baseline[a] | |
|---|---|---|
| Patient Group | Day 8 | Day 84 |
| Wild type, high dose | 8.2 | 18.5 |
| Mutant, high dose | 1.8 | 5 |
| Placebo | -0.7 | -1.4 |

[a]Data extrapolated from published data

Approximately 6% of asthma patients do not carry a wild-type allele at the 5-LO core promoter locus, and this data indicates that these patients would not benefit from 5-LO inhibitor drug therapy. Further, these data indicate that there is evidence to reasonably identify patients, i.e. stratification based upon 5-LO genotype, to appropriately treat patients with asthma.

A recent double blind, placebo controlled crossover designed pharmacogenetic restrospective clinical trial of a β2-adrenoreceptor polymorphism was implemented to analyze the significance of β2-adrenoreceptor polymorphisms (Tan et al. Lancet 350:995-999). *In vitro* studies have suggested that polymorphism of the β2-adrenergic receptor may influence the desensitization induced by β2 agonists. Twenty two moderately severe asthmatics were enrolled into a placebo controlled cross-over study of formeterol (a β2-adrenergic agonist). The patients were divided into groups by allelic variances: 1) at codon 16, homozygous arginine (n=4), heterozygous arginine/glycine (n=8), and homozygous glycine (n=10); and 2) at codon 27, homozygous glutamine (n=5), heterozygous glutatmine/glutamic acid (n=11), and homozygous glutamic acid (n=6). Genotypic analysis determined that individuals who were homozygous for glycine at codon 16 were also homozygous for glutamic acid at codon 27. The results were as follows:

| | Degree of Brochodilator Desensitisation after Formoterol Therapy[1] | | | |
|---|---|---|---|---|
| Polymorphisms of the β2-adrenergic | 6 hour $FEV_1$ | Maxim al $FEV_1$ | 6 Hour $FEF_{25-75}$ | Maxim al |

| receptor | | | | $FEF_{25-75}$ |
|---|---|---|---|---|
| Gly 16 (n=10) | 80% | 48% | 103% | 73% |
| Arg 16 (n=4) | 28% | -8% | 23% | -35% |
| Gly/Arg16 (n=8) | 57% | 48% | 70% | 50% |
| | | | | |
| Glu27 (n=6)[2] | 73% | 35% | 90% | 68% |
| Gln27 (n=5) | 47% | 3% | 38% | -15% |
| Glu/Gln27 (n=11) | 65% | 52% | 70% | 45% |

[1]Data extrapolated from published graphs.

[2]All individuals homozygous for Glu27 were also homozygous for Gly16.

The homozygous glycine at position 16 was associated with individuals who were prone to bronchodilator desensitization than at arginine at position 16: the mean $FEV_1$ desensitisation was 80% for Gly16 homozygotes versus 28% for the Arg16 homozygotes. Similar results were observed for the 6 hour $FEV_1$ and the FEF.

For the polymorphism at codon 27, the mean for the Glu27 homozygous individuals demonstrated greater desensitization than those who were homozygous for Gln27.

The allelic variance, glycine at position 16 appeared to dominate over the putative protective effects of the mutation of glutamic acid at position 27.

The effects of the codon 16 and 27 polymorphism in the β2-adrenoreceptor on β2-agonist desensitization, as observed in the above data, suggest that there may be an identifiable subset of patients for whom β2-adrenergic receptor desensitization occurs in the presence of long-acting or repeated use of β2-agonists.

Thus, one skilled in the art, will be able to utilize the presently described pharmacogenetic techniques to identify the allelic variances with the coding region of the β -adrenergic receptor or other receptor proteins that are similar to the β -adrenergic receptor, including but not limited to those variances for those genes listed in Tables 4, 15, and 21 and those 7-membrane spanning receptor G-protein coupled receptors. In this way, a skilled practitioner will be able to utilize the methods, protocols, and techniques that are described in the detailed description and those known in the art to identify the gene targets, allelic variance or variances, and candidate drugs that affect these pathways. Further, one can design and implement a strategy that incorporates a diagnostic test to genotype the individual for a given allele or alleles or halpotype, grouping these candidates by genotype, and testing a β-adrenergic agonist or other candidate therapeutic product for the affect of the pharmacogenomic difference between or among the groups.

As described above, there is evidence to suggest that there are safety response differences to drug therapy in asthma which may be attributable to genotypic differences between individuals. There is provided in this invention examples of gene pathways that are implicated in the disease process or its therapy and those that potentially cause this variability. The Detailed Description above demonstrates how identification of a candidate gene or genes and gene pathways, stratification, clinical trial design, and implementation of genotyping for appropriate medical management of a given disease can be used to identify the genetic cause of variations in clinical response to therapy, new diagnostic tests, new therapeutic approaches for treating this disorder, and new pharmacuetical products or formulations for therapy. Gene pathways including, but not limited to, those that are outlined in the gene pathway Tables 1-6, preferably Table 4, and pathway matrix Table 9 and discussed below are candidates for the genetic analysis and product development using the methods described above.

As an example of identification of the primary gene and relevant polymorphic variance that directly affects efficacy, safety, or both one could select an gene pathway as described in the Detailed Description, and determine the effect of genetic polymorphism and therapy efficacy, safety, or both within that given pathway. For example, referring to Table 9, genes involved in cytokine-mediated immune regulation, non-cytokine mediated immune regulation (including, but not excluded to, cyclophilins, corticosteroids), cell mediated inflammation involving apoptosis, adhesion and migration, protease and protease inhibitors, complement, degranulation (platelets, mast cells, neutrophils, ans eosinophils), release of inflammatory modulators (including membrane lipids, prostaglandin, platelet activating factor, leukotrienes, histamine, nitric oxide), vascularization mediators (including endothelin and vascular endothelial cell growth factor), neurotransmitters and peptide hormone inflammation modulators (including adrenergic, purinergic, cholinergic, ion channels, tachykinin, neurokinin, substance P, bradykinin, parathyroid hormone, melanocortin and adrenocorticotrophic hormones, and modulators of general cell growth pathways the optimization of therapy of by an agent can be achieved by determining whether the patient has a predisposing genotype in which the selected agents are more effective and or are more safe. In considering an optimization protocol, one could potentially predetermine the genotypic profile of these genes involved in the manifestation of the adverse effect, or those genes preeminently responsible for drug response. By embarking on the previously described gene pathway approach, it is technical feasibility to determine the relevant genes within such a targeted drug development program.

*Description of Mechanism of Action Hypotheses for Future Drug Development*

There are many potential mechanisms that may serve as targets for candidate therapeutic interventions. For example, phosphodiesterase inhibitors to PDE4; T-lymphocyte-eosinophil interactions inhibition: targeting the factors involved in the regulation of the TH2(CD+4) differentiation and/or activation by soluble factors (cytokines (IL-4, IL-5); co-stimulatory molecules (B7-2/CD86); and transcritpion factors (GATA-3, AP-1). These targets may be available to limit the TH2 cell involvement in the initiation of asthmatic inflammation.

Suppression of eosinophil adhesion with consequent inhibition of influx into the lung is a strategy to suppress asthmatic airway inflammation. Such inhibition may be mediated through inhibitors directed towards very late antigen-4 (VLA-4), monoclonal antibodies directed towards VLA-4, intracellular adhesion molecule 1 (ICAM1), and alpha 1,3-fucosyltransferase VII (an enzyme which regulates sekectin function). Furthermore, molecules may be targeted to suppress the expression of adhesion molecules (e-selectin, vascular cell-adhesion molecule 1 (VCAM-1), and ICAM1).

There are a group of chemokines that contain a cysteine-X-cysteine motif, such as IL-8 that are effectors of acute inflammatory episodes, whereas cysteine-cysteine chemokines, such as macrophage inhibitory peptide 1 (MIP-1), eotaxin, RANTES, or macrophage chemotactic peptide 1 (MCP-1) act as chronic mediators of inflammation. These molecules may be appropriate targets for inhibiting either the acute or chronic inflammatory pathway.

Cysteinyl leukotrienes have a central role in the development of chronic asthma, and antagonists (i.e., $CysLT_1$) may be able to ablate the actions of this ligand. These novel leukotriene receptor agonists may have potential for anti-inflammatory effects. Endothelin receptors may also be a target, with endothelin antagonists to specific receptor subtypes $ET_A$ or $ET_B$. Other receptors known to be involved in the inflammatory process that may be potential targets are the tachykinin NK1 receptors and selective ligands to the NK1/NK2 receptors.

Induction of cyclooxygenase and the consequent increase in prostaglandin release is associated with the development of inflammation. Antisense oligonucleotides directed against the receptor types NK-kB, major basic protein, 5-lipoxygenase, leukotriene C4(LTC4 synthetase, IL-4, IL-5, IL-8 and adenosine have been developed that are inhalable products that can directly block the expression of these mediators of the inflammatory response.

Other areas of drug target development include immunobiology of the airways i.e., TH1 and TH2 and their involvement in the immune response, synthesis of immunoglobulin, IgE, integrins, inhibition of αIL5 and αIL5 monoclonal

antibody, soluble IL4 receptor, neurokinin receptor antagonist, chemokine inhibitors.

The inflammatory response is also being evaluated in terms of the effects of NO2, SO2, and ozone on the subsequent effect on airway response to these potential allergens. As well as adhesion molecule expression, cytokine production, and cytokine gene transcription factors.

Optimization of nonsteroidal or steroidal antiinflammatory agents , or agents aimed at a mechanism of therapy of the underlying etiology of asthma further demonstrates the utility of selection of a potential asthma patient that has a predisposing genotype in which selective antiasthmatic or other agents may be more effective and or have an more desirable safety profile. In considering an optimization protocol, one could potentially predetermine variance or variances within the nonsteroidal antiinflammatory pathway, steroid antiinflammatory pathway, or antiinflammatory mediated intracellular mechanism of action that is preeminently responsible for antiasthmatic drug response. By embarking on the previously described gene pathway approach, it is technically feasible to determine the relevant genes within such a targeted drug development program for asthma.

A sample of therapies approved or in development for preventing or treating the progression of asthma currently known in the art is shown in Table 47. In this table, the candidate therapeutics were sorted and listed by mechanism of action. Further, the product name, the pharmacologic mechanism of action, chemical name (if specified), and the indication is listed as well.

**Example 25**

Inflammatory Bowel Disease

*Description of Inflammatory Bowel Disease*

Inflammatory bowel disease (IBD) is a broad clinical term that includes idiopathic chronic inflammatory bowel diseases including Crohn's disease (CD) and ulcerative colitis (UC) which can be distinguished from inflammatory bowel disease of known origin including diverticulitis, radiation enteritis, colitits, drug or toxin-induced enterocolitis, or vasiculitis of the intestinal tract. UC is a term that encompasses a broad category of diffuse, continuous, and superficial inflammation of the colon, which begins within the rectum and extends proximally. The condition is limited to the colon and large intestine, with limited involvement of the small intestine. In UC, the inflammation primarily affects the mucosal process and is not transluminal within these anatomical regions. CD is characterized by focal, asymetric, transmural inflammation affecting any portion of the gastrintestinal tract, i.e. from the mouth to the anus. The focal localization and possible extent of the

inflammation distinguishes UC from CD. There are currently approximately 35-100 and 10-100 CD per 100,000 Americans diagnosed with UC or CD, respectively.

Clinically, patients with UC experience variable stool consistencies from constipation to diarrhea, low-grade fever, malaise, nausea, vomiting associated with defecation, night sweats, arthalgias, dehydration, tachycardia, and symptoms of abdominal tenderness. There can be rectal bleeding, tenemus, and passage of mucopus.

Patients with Crohn's disease experience symptoms of peptic ulcer disease, nausea, vomiting, and epigastric pain. Transmural inflammation leads to fibrosis and transluminal narrowing. In some cases, the narrowing leads to signs and symptoms of intestinal obstruction including nausea, vomiting, waves of abdominal pain, and a reduced output of stool. Patients with colonic CD are likely to experience abdominal pain, cramping or localized pain, rectal bleeding, and diarrhea. Weight loss is common among CD patients due to malabsorption of nutrients and reduced food intake due to minimization of postprandial symptoms.

There are extraintestinal manifestations of inflammatory bowel disease affecting the following processes including: nutritional and metabolic abnormalities, hematologic abnormalities, skin and mucous membranes, musculoskeletal, hepatic and biliary abnormalities, renal complications, and optic complications. These complications are associated when the colon or intestinal tract is inflammed. These complications are clinically menifset as joint swelling or pain, erythema nodosum, pyoderma gangenosum, sclerosing cholangitis, conjunctivitis, or uveitis.

There is an increased risk for the development of gastrointestinal cancer in patients with IBD. In both UC and CD, there is an increased risk of adenocarcinoma of the intestine. This is not correlated to the intensity of the first attack, subsequent course, or and specific medical therapeutic approach. Therefore routine screenig for dysplasia and neoplasia is warranted.

*Current Therapy of Inflammatory Bowel Disease*

Strategies for the therapy of inflammatory bowel disease includes antiinflammatory agents, and immunomodulation.

Antiinflammatory agents include the use of glucocorticoids and the aminosalicylates. Glucocorticoids act by modulation of the immune response. Corticosteroids affect the inflammation within the gastrointestinal tract by decreasing growth and development of mast cells, inducing apoptosis, suppressing lymphocyte generation of IL-5 and other cytokines, inhibiting some mediator release, inhibiting cytokine production, inhibiting the transcription of cytokines (for example IL-8, TNF-$\alpha$, prototypic antiviral chemokine (regulated-on-activation

normal T-expressed and secreted, RANTES), and GM-CSF), and inhibiting nitric oxide synthesis.

5-aminosalicyclic acid (5ASA) is a salicylate that is used for the treatment of IBD, is not orally active, is poorly absorbed and is inactivated by intestinal bacteria, and is delivered as a suppository or rectal suspension enema. Oral formulations can be used to deliver active drug to the lower intestine which are cogeners of 5ASA. The aminosalicylates are potent antiinflammatory agents that inhibit cyclooxygenase (COX), arate limiting enzyme in the protaglandin and leukotriene pathway.

Immunosuppressive agents are also used to modulate the inflammatory/immune response. There are four broad categories of immunosuppresive agents that have distinct mechanisms of action: inhibition of ribonucleotide synthesis which acts to inhibit the proliferation of T-cell clones (6-mercaptopurine), inhibition of folic acid which acts to inhibit T-cell and B-cell function as well as decrease IL-1 and IL-6 activity (methotrexate), inhibition of T-cell receptor stimulated transcription of lymphokine genes which act to ihibit the production of IL-2 and IL-2 receptors as well as inhibit certain cytokines (TNF-α, IFN-γ) (cyclosporin and FK506), and inhibition of guanosine nucleotide synthesis which acts as cytostatic effects on lymphocytes (mycophenolate). Each of these catgoiies of agents have been employed for the therapy of IBD.

Recently a chimeric monoclonal antibody was approved for use in the treatment of moderately to severe active Crohn's disease for those patients that are unresponsive to conventional therapy. This monoclonal antibody is specific for TNF-α and can remove TNF from the bloodstream before it reaches the site of inflammation.

Crohn's disease may progress to a level and extent in which surgical removal of the localized inflammation is warranted. Surgery has been indicated for recurrent intestianl obstruction, complicated fistulas, intractable hemorrhage, disease refractory to medical therapy, growth retardation refractory to therapies, or cancer. The surgical procedures vary from excision of a localized, diseased portion of the gastrointestinal tract to removal of large portions, i.e. the entire colon (colectomy). Surgical excision of the inflammed region or to correct complications such as blockage, perforation, abscess, or bleeding can result in a substantial relief of symptoms.

*Limitations to Current Therapies for IBD*

Salicylate associated side effects include dyspepsia, gastric or small bowel bleeding, ulceration, renal insufficiency, confusion, rash, headache, hepatic toxicity. NSAIDs also reversibly inhibit platelet aggregation and prolong bleeding time.

Glucocorticoid associated side effects include increased appetite, weight gain, fluid retention, acne, ecchymosis, development of cushoid facies, hypertension, hyperkalemia, diabetes, hyperglycemia, hyperosmolar state, hyperlipidemia, hepatic steatosis, atherosclerosis, myopathy, aseptic necrosis, osteoporosis, ulcers, pancreatitis, psuedotumor cerebri, pyschosis, glaucoma, cataract formation, vascular necrosis, increased suseptibility to infection, impairment of the hypothalamus-pituitary axis, decreased thyroid hormone serum binding protiens, and impaired wound healing.

Agents involved in immunomodulation have the following undesirable side effects including antimetabolites: hepatic compromise including hepatic fibrosis, ascites, esopageal varices, cirrhosis, pneumonitis, myelosuppression; immunosuppressives: myelosuppression, (cyclosporine: renal insuffienciency anemia, hypertension.

Monoclonal antibody to TNF proteins therapies have been shown to generate a human-antimouse antibody response (HAMA). However, patients on immunosuppressive agents such as glucocorticoids and others are less likely to generate antibodies to the treatment antibody. Delayed hypersensitivty is demonstrable 2 to 4 years after initial treatment in 25% of the patients treated with the chimeric antibody. Further, there are patients that develop a serum sickness reaction which includes fever, and joint swelling that requiring hospital admission. A positive antinuclear antibody (ANA) occurred in 24-36% of the patients analyzed. Nine percent of the patients developed anti-DNA antibodies, less than 1% developed a lupus-like reaction requiring steroid therapy.

In surgical therapy of IBD, reccurring inflammation and relapse, after excision procedures occurs in 75% of the patients. Attempts have been made to include salicylate therapy after resective surgery, however, the inflammation reccurrance rate in that group was 52%.

*Impact of Stratification Based Upon Genotype in Drug Development for Drugs, Compounds, or Candidate Therapeutic Interventions for Autoimmune Disease*

Thiopurine methyltransferase (TPMT)

The thiopurine S-methyltransferase (TPMT) is a cytosolic enzyme whose precise physiological role is unknown. This enzyme catalyzes the S-methylation of widely used immunosuppressive or cytotoxic thiopurine drugs such as 6-thioguanine, 6-mercaptopurine and azathioprine.8 The *in vivo* activity of this cytosolic enzyme is characterized by interindividual and interethnic variability caused by the genetic polymorphism of the TPMT gene, which was discovered, using pharmacogenetic techniques, by the existence of three major phenotypes, high

(HM), intermediate (IM) and deficient (DM) methylation. As a consequence, individuals greatly differ in detoxication of thiopurine drugs to 6-methylmercaptopurine as well as the occurrence of side effects or therapeutic efficacy. Using genomic techniques, PCR-SSCP (polymerase chain reaction – single strand conformation polymorphism), Spire-Vayron de la Moureyre et al. 9 have defined the mutational and allelic spectrum of TPMT in a group of 191 Europeans. In this analysis, PCR-SSCP techniques identified allelic variances in the entire coding sequence, the exon-intron boundaries, the promoter region and the 3'-flanking region of the genes. Six mutations were detected throughout the ten exons and seven TPMT alleles were characterized. Within the promoter region, six alleles corresponding to a variable number of repeats (VNTR) were identified. The TPMT phenotype was correctly predicted by genotyping for 87% of individuals. A clear negative correlation between the total number of repeats from both alleles and the TPMT activity level was observed, indicating that VNTRs contribute to inter-individual variations of TPMT activity. This VNTR polymorphism can be considered responsible for shifts to lower or higher TPMT activities observed among discordant individuals. Seven out of the nine phenotyped HMs but genotyped IMs were carrier of a total of eight VNTR repeats. This low number of repeat can account for the switch to high TMPT activities of these samples.

One in 300 patients with IBD are homozygous-deficient for TPMT. The clinical relevance for this deficiency is that TPMT is the enzyme responsible for the conversion of 6-MP to 6-MMP, and the AZA compounds to 6-TG. In TPMT deficient patients, higher levels of 6-TG and 6-MMP are then produced and are associated with significant leukopenia. In general, patients produce variable levels of 6-TG and 6-MMP as determined by their intrinsic enzyme systems. Higher 6-TG levels are correlated with good therapuetic response, but produce leukopenia. Higher 6-MMP levels correlate with hepatotoxicity and in recent studies with leukopenia.

There is evidence to suggest that there are safety response differences to drug therapy in IBD which may be attributable to genotypic differences between individuals, one example being the TPMT gene described above. There is provided in this invention examples of other gene pathways that are implicated in the disease process or its therapy and those that potentially cause this variability. The Detailed Description above demonstrates how identification of a candidate gene or genes and gene pathways, stratification, clinical trial design, and implementation of genotyping for appropriate medical management of a given disease can be used to identify the genetic cause of variations in clinical response to therapy, new diagnostic tests, new therapeutic approaches for treating this disorder, and new pharmacuetical products

or formulations for therapy. Gene pathways including, but not limited to, those that are outlined in the gene pathway Tables 1-6, preferably Table 4, and Table 9 and discussed below are candidates for the genetic analysis and product development using the methods described above.

*V. Description of Mechanism of Action Hypotheses for Future Drug Development*

The majority of the hypotheses for future therapuetic interventions for inflammatory bowel disease are based upon the understanding the immunologic mechanisms that cause and perpetuate the inflammation within the gastrointestinal tract. Although the initiating event is elusive, the resulting immunologic events have been studied. All of the gastrointestinal enterocytes have immunologic function. Under physiologic conditions, these enterocytes selectively activate CD8+ nonspecific suppressor cells, in response to inflammation. In patients with IBD, these enterocytes selectively stimulate the development of CD4+ helper T cells which can respond in two ways 1) the Th1 response which involves the activation of IL-2 and IFN-g and leads to delayed hypersensitivity and cellular immunity and 2) the Th2 response which involves IL-4, IL-5, IL-6, and IL-10 and leads to antibody response and humoral immunity. Both Th1 and Th2 responses are genetically controlled and are coordinately regulated, i.e. Th1 response stimulation results in down regulation of Th2 response and vice versa. It has been demonstrated that in UC patients the Th2 response is favored and in CD patients the Th1 response is favored.

A humanized (95% human, 5% mouse) version of the chimeric antibody (75% human, 25% mouse) to TNF is currently under development. Some antiidiotypic antibodies are generated, but it doesn't appear to stimulate a delayed hypersensitivity, no stimulation of anti-DNA antibodies, or lupus-like reactions.

Mediators of the immune response including intracellular adhesion molecule (ICAM-1) inhibitors (antisense molecules or others), IL-10, IL-11 have been tested in humans. Further, and anti-CD4 monoclonal antibody which has been shown to interfere with the interaction of the CD4 molecule and the HLA class II molecules leading to an inhibition of antigen presentation has been tested.

Thalidomide (inhibitor of TNF, acceleration of the degradation of the TNF mRNA) is also under consideration.

It has been noted that individuals who smoke tobacco products have a lower incidence of IBD. Therefore, understanding the immune response and correlation with nicotinic chloinergic pathways is under investigation.

A sample of therapies approved or in development for preventing or treating the progression of IBD currently known in the art is shown in Table 48. In this table, the candidate therapeutics were sorted and listed by mechanism of action. Further, the product name, the pharmacologic mechanism of action, chemical name (if specified), and the indication is listed as well.

**Example**

Hepatitis C

Selecting Optimal Therapy for HCV Patients

Genetically Determined Variation in Response to Interferon α

Treatment of hepatitis C virus (HCV) infection with interferon α is expensive, benefits a minority of patients, and produces side effects in a significant fraction of patients. Addition of ribavirin increases efficacy, but combination therapy remains expensive and still falls well short of providing a lasting benefit to most patients. It would therefore be desirable to identify prospectively those patients likely to have a sustained response to treatment. Ideally a diagnostic test would also predict what dose of interferon and ribavirin, administered for what length of time, will afford to each patient the best chance of a sustained response. Pre-treatment identification of patients likely to suffer serious toxic side effects would also be desirable.

The best characterized predictors of response to interferon α therapy are viral load and HCV genotype. Low viral load before therapy is predictive of a positive response. However, demonstration of decreased viral load after initiation of therapy is currently the best predictor of response to therapy. There is no consensus on the optimal time after initation of therapy for measuring viral levels; periods ranging from 2 weeks to four months have been proposed. The viral load test is not very effective at discriminating long term responders from those patients who suffer rebound of HCV infection within 6 months after treatment. Also, the ideal test would be performed in advance of any treatment, thereby saving the considerable costs associated with even short term therapy. In search of other predictive indices, over 100 controlled clinical studies have examined a variety of viral and host factors in responders and nonresponders. Genetic variation in both HCV and host genes has been shown to independently influence patient response to interferon α treatment. A consensus has emerged regarding the interaction of viral genotype and treatment response, however the contribution of host factors to treatment response has not been as well investigated. There are a number of promising recent findings suggesting that polymorphisms in regulators of human immune function are correlated with response to interferon α.

*Viral genome variation*

Comparison of sequenced HCV genomes reveals considerable variation in viral sequence, with at least 6 major types and well over a dozen minor types recognized. The geographical distribution of viral types is nonrandom, perhaps accounting for some of the apparent racial heterogeneity in the natural history of HCV infection. HCV is present in each patient as a heterogeneous population of viral quasispecies, with the degree of heterogeneity differing among patients. Despite these complexities, there are strong correlations between predominant viral type and treatment response. In general, patients with genotype 1 (especially 1b) respond poorly to interferon α, with many studies showing response rates under 10 percent. Patients with genotype 2 or 3 do well, with response rates typically greater than 40 percent. Most viral genotyping is based on a short variable segment, however there are multiple segments of the viral genome that vary, and some studies have found that more detailed viral genotyping, for example of the 5' untranslated region, provides stronger correlations with treatment response.

*Human genome variation*

A recent study suggests that there is significant variation in response to interferon α treatment among racial groups in the US, even after controlling for the effect of different HCV types. This finding suggests that host genetic variation may be an important factor in response. A number of candidate genes have recently been tested for correlation with interferon α response. The best studied genes are regulators of immune function such as IL-6, IL-10 and TNFα. One study, for example, found that patients with high expression of IL-10 (attributable to a specific haplotype) tend to respond poorly to interferon, perhaps due to impaired immune response. IL-4, IL-12 and TGF-β levels have been correlated with treatment response in some (but not all) studies, however no genetic analysis has been performed. Similarly, hepatic levels of interferon α–β receptor have been correlated with response to interferon, but no genetic analysis has been performed to determine whether polymorphism affects receptor levels. HLA alleles have also been correlated with response to interferon, particularly the A24-B54-DR2 haplotype. A number of other compelling candidate genes have not been investigated. For example, a recent report shows that HCV can enter cells via the low density lipoprotein receptor. If so, the well studied amino acid polymorphisms of the LDL-R should be investigated for effects on disease course and response to treatment. There are also likely to be genetic factors that influence response to ribavirin; for example, the drug must be transported across the plasma membrane and then

phosphorylated before becoming a substrate for viral enzymes. The transporters and kinases responsible for these processes may be worth genetic investigation.

An optimal test for selecting treatment for HCV infection would (i) suggest the optimal therapeutic regimen (interferon alone, interferon and ribavirin, or some other combination), (ii) suggest the optimal dose and duration of treatment, (iii) predict sustained responders vs. short term responders, and (iv) predict patients likely to suffer serious adverse effects. At least three areas should be further investigated to better predict the reponse to interferon α treatment. First, it is not clear that conventional viral genotyping methods, focusing on the 5' untranslated region, capture all of the aspects of viral sequence variation that affect viral biology. Additional genetic determinants of viral pathogenicity should be investigated. Second, the human gene variants that have been associated with response need to be more thoroughly investigated, and interactions between human candidate gene alleles, as well as perhaps between human genes and viral genes, should be tested. Third, recent work suggests a number of new host proteins that may affect response to interferon, and proteins that mediate response to ribavirin have not yet been investigated. The genes encoding these proteins should be thoroughly investigated. With additional information on candidate genes available it should be possible to construct a plan, ideally via retrospective analysis of clinical trial data, for first assesing the impact of variation in each of the candidate genes, then examining gene x gene interactions, and finally reducing the number of tests to a much smaller number for confirmatory prospective trials.

In Table 49, there a list of the candidate therapeutic interventions that in development for Hepatitis. One skilled in the art could apply, as described in the text, the methods of this invention to ascertain whether there is a gene in the inflammatory pathway that may be involved in the efficacy, safety, or toxicities of these candidate interventions.

**Example 27**

Pro12Ala Substitution in PPARγ2 Affects Insulin Sensitivity

Peroxisome proliferator-activated receptors (PPARs) are members of the nuclear hormone receptor family of DNA binding transcription factors. PPARs form heterodimers with retinoid X receptors and the resultant heterodimers, in coordination with coactivators and corepressors,bind to DNA and activate transcription of various genes. The PPAR superfamily includes receptors that mediate the size and number of peroxisomes in response to a diverse group of chemicals both naturally occurring and xenobiotics. Endogenous ligands thought to activate the PPARs are arachidonic acid, oleic acids, andogenous molecules (fatty

acids or steroids), C18 unsaturated fatty acids, peroxisome proliferation activators, and others (see Table 5). Diverse chemicals can activate the PPARs: herbicides, leukotriene antagonists, plasticizers (phthalate ester plasticizers used in the production of vinyl plastics), the fibrate class of hypolipidemic agents, thiazolidinediones. Overstimulation of these receptors can result in heptomegaly, liver hyperplasia, and possibly heptocellular carcinoma. There are three known PPARs, α, γ, δ. PPARα is believed to be involved in the regulation and control of fatty acid oxidation enzymes. PPARα is has been shown to have high expression rates in heart, adipose, and liver. PPARγ is believed to be involved in adipocyte differentiation. PPARγ is expressed in high levels in adipocyte tissues. PPARδ (NUC1) is believed to be involved in a family of DNA binding proteins that are involved in adipogenesis and may be involved in early development. PPARδ has been has been shown to have high expression in heart, kidney, and lung.

PPARα is involved in the metabolic control of the expression of genes encoding fatty acid oxidation enzymes. Data from several experimental strategies have supported the hypothesis of the mechanism of action of PPARa: 1) PPARa is necessary for the induction of peroxisomal biogenesis in response to peroxisomal proliferating agents; 2) the target genes of PPARs are enzymes involved in cellular fatty acid oxidation which include mitochondrial, peroxisomal, and cytochrome P450 pathways; 3) PPARα is activated by fatty acids or inhibitors of mitochondrial long-chain fatty acid import. It has been shown that PPARα modulates the expression of genes encoding lipid metabolism enzymes, lipid transporters, or apolipoproteins. In an animal model of hyperlipidemia, activators of PPARα was shown to decrease the lipid production in hepatocytes, however PPARα activation also demonstrated tumor promotion within the same animals. Ligands that can specifically activate the lipostat enzymes while not turning on tumor production would be advantageous.

PPARγ is thought to be involved in the differentiation of preadipocytes to adipocytes. Overexpression of PPARg in a non-adipose cell, i.e. nonadipogenic fibroblasts, results in the conversion to fat-laden adipocyte-like cells after exposure to a PPARγ ligand. Another transcription factor family involved in adipogenesis is the CAAT/enhance binding protein, CEBP. CEBPa is expressed in high abundance in adipose tissue and may play a direct role in establishing and maintaining the fully differentiated adipocyte phenotype. This hypothesis is based mainly on the data that indicates CEBPα is expressed late in adipogenesis and after key enzymes are induced. In other studies it has been shown that PPARγ and CEBPα expression can both be induced by CEBPβ and CEBPδ. PPARγ and CEBPα both induce the

expression of each other as well as activate and maintain the adipocyte proliferative and growth differentiation program.

The PPARγ gene has two transcription start sites and translation results in two distinct proteins PPARγ1 and PPARγ2. Both are highly expressed in adipose tissue. As in other nuclear hormone receptors, PPARγ is dependent on ligand activation. Currently, known biological ligands are 15-deoxy-$\Delta^{12,14}$ prostaglandin J, other prostanoids, and products from the linoleic acid pathway, i.e. oxLDL, HETE, 13-HODE and 9-HODE. Xenobiotics from the thiazolidinedione group, i.e. troglitazone, ciglitazone, and pioglitazone can directly activate PPARγ.

Modulation of PPR activities are thought to be effective strategies for the development of products for therapy of cancers (breast, prostate, and acute promyleocytic leukemia), metabolic diseases including thyroid disease and diabetes mellitus. PPARγ is expressed at significant levels in human primary and metastatic breast adenocarcinomas. Experimental evidence has suggested that the PPARγ transcriptional pathway can induce terminal differentiation of malignant breast epithelial cells. Ligands known to activate PPARγ have been shown to cause lipid accumulation, reduction of growth rate, and a reversion to a differentiated, less malignant state in studies of cultured breast cancer cells. Further, inhibition of MAP kinase, a negative regulator of PPARg, enhances the activation by a PPARγ ligand (i.e. thiazolidinedione) sensitivity.

In studies of an animal model of diabetes, ligands that specifically activate PPARγ (i.e. troglitazone), normalization of elevated glucose levels in obese animals was demonstrated. Studies have been conducted to ascertain the efficacy of thiazolidinediones to treat NIDDM. One product troglitazone (Rizulin) has achieved approval for human therapeutic use in the U.S.

In a recent study, it was determined whether genetic variation in the PPARγ coding region was associated with obesity and insulin sensitivity or resistance, as well as type II diabetes mellitus (Deeb et al., Nature Genetics, 1998). It was determined that there was a single polynucleotide base substitution (C/G) which lead to a substitution in the coding sequence of proline to alanine at amino acid position 12 (Pro12Ala). The study included two human populations, Finns and Japanese-Americans. It was determined that the relative frequency of the the alanine alle frequency in the Finn study population (nondiabetic, including some with impaired glucose tolerance) was 0.12 whereas the Japanese-American frequency was 0.022 in type II DM patients, 0.039 in patients with impaired glucose tolerance, and 0.093 in normal subjects. In both populations there was an association of the Ala allele frequency and lower fasting insulin levels and body mass index, as well as higher insulin sensitivity; in the Finn population the values achieved statistical significance.

The study further demonstrated a functional correlation of the population data with in vitro PPARγ transcription factor binding affinities. In these experiments, it was shown that the Ala-isoform demonstrated a two- to five-fold decrease in relative affinity for the identified peroxisome proliferator response element, as well as a 36% faster off rate in comparison with the values detected for the PPARγ Pro isoform. Confirmatory data in the form of reduced detectable transactivation by the PPARγ ligand in the case of the PPARγ Ala isoform.

In addition, there is data to suggest that PPARγ mRNA expression levels are reduced in obese individuals and that the the ratio of mRNA encoding PPARγ is positively correlated with body mass index.

These data suggest that there is an association of reduced trasnscription activation by the Ala PPARγ allelic variant. Further suggesting, that there is a molecular mechanism for the observed body mass index and insulin sensitivity in the individuals having these allele polymorphisms. The data reported suggests that via reduced transcription of target genes that are involved in regulation of glucose homeostasis.

**Example 28**

Sulfonylurea receptor silent polymorphism and insulin levels

In a sub-population that is approximately three times more likely to acquire type II DM, Mexican-Americans have higher insulin concentrations, and are more likely to exhibit insulin resistance. It has further been determined that in this population of Americans, hypeinsulinemia is a risk factor for the development of type II DM.

The high affinity sulfonylurea receptor (SUR1) is known to be involved in the reguilation of insulin secretion. This receptor may be involved in type II DM. The SUR1 gene product is a functional part of the pancreatic β-cell KATP ion channel. The channel complex is composed of a two subunits, the sulfonylurea binding domain and the b-cell KATP channel responsible for conducting an inward rectifying potassium current. With the β-cell, metabolism of glucose produces changes in the relative concentrations of ADP and ATP which leads to a reduction of the KATP channel activation, causing a depolarization of the β-cell membrane and exocytosis of insulin. Within the nucleotide-binding fold region (NBF) of the SUR1, mutations have been shown to be autosomal recessive and lead to clinical familial hyper insulinemia. Other mutations in the SUR1 have been associated with Beckwith-Waldemann-syndrome associated malignant insulinomas.

In exon 31, there is a silent polymorphism (AGG/AGA) that encodes an arginine residue at position 1272. In the Mexican American study population that had the AGA genotype, there were higher fasting and 2 hr. insulin levels as well as a higher proinsulin to insulin ratio than those observed in the wild-type genotype sub-group. Between the two groups there were similar values for fasting glucose, body mass index, and waist circumference measurements.

| Test Parameter | AGA | AGG | P value |
|---|---|---|---|
| Fasting insulin* | 113.4 | 82.8 | 0.043 |
| 2 hr. insulin* | 849.6 | 498.6 | 0.0003 |
| Proinsulin/insulin ratio | 0.068 | 0.056 | 0.030 |

*values in pmol/l

These data taken togethere suggests that there is an association between the SUR1 allelic variant and hyperinsulinemia in normal individuals from a high DM risk ethnic group.

**Example 29**

Vitamin D Receptor and Estrogen Receptor Polymorphisms and Response to Hormone Replacement Therapy

Bone mineral density (BMD), a predictor of risk of bone fractures, decreases rapidly in postmenopausal women. Hormore replacement therapy (estrogen) reduces the rate of or prevents the decrease in BMD. Genetic factors contribute to 60-80% of BMD variation. In a recent study (Deng et al., Hum Genet 103:576-585), it was shown that hormone receptor polymorphisms affect BMD in elderly women

and that genotype should be considered when prescribing hormone relacement therapy (HRT) to preserve bone mass in elderly Caucasian women.

A population of 108 women participated in the study. They were genotyped for polymorphic differences in their vitamin D (VDR) and estrogen (ER) receptors. Using restriction endonuclease specific sites within these genes, it was determined that the VDR has a polymorphic *Bsm*I site (B or b) and the ER has two polymorphic sites, *Xba*I (X or x) and *Pvu*II (P or p). In the placebo and HRT groups, the VDR and ER genotype groups had significant affect on the BMD measurements. An analysis of the gene-by-gene interaction revealed that the level of significance was reduced. The amount of variation in BMD attributable to the VDR and ER polymorphisms varied from approximately 1% (for the total body bone mineral content changes in the placebo or HRT groups) to approximately 18.7% (for the spine bone mineral density changes occuring in the HRT group). Significant genotype effects were observed in the xx, PP, or bb groups having a larger decrease of bone mass during the study period, whereas a genotype of XX, pp, or BB is associated with smaller decreases (or larger increase) of bone mass.

This study demonstrates and interaction of drug response with genotype with age/reproductive status.

**Example 30**

Cholesterol ethyl-transferase (CETP)

A well studied polymorphism in the first intron of the gene encoding cholesterol ester transfer protein (CETP) provides an example of a polymorphism in the non-coding region of a gene that has with an impact on drug efficacy via a recessive genetic mechanism.

The high-density lipoprotein (HDL) cholesterol concentration is inversely related to the risk of coronary artery disease. CETP has a central role in the metabolism of HDL and may therefore alter the susceptibility to atherosclerosis. The DNA of 807 men with angiographically documented coronary atherosclerosis was analyzed for the presence of a polymorphism in the gene coding for CETP. The presence of a DNA variation in a Taq I restriction enzyme site was referred to as B1, and its absence as B2. All patients participated in a cholesterol-lowering trial of the drug pravastatin designed to reduce cholesterol synthesis by inhibiting HMGCoA Reductase, and thereby arrest progression of, or induce the regression of coronary atherosclerosis and were randomly assigned to treatment with either pravastatin or placebo for two years. The B1 variant of the CETP gene was associated with both higher plasma CETP concentrations (mean [±SD], 2.29±0.62 μg per milliliter for the B1B1 genotype vs. 1.76±0.51 μg per milliliter for the B2B2 genotype) and lower

HDL cholesterol concentrations (34±8 vs. 39±10 mg per deciliter). In addition, a significant dose-dependent association between CETP genotype and the progression of coronary atherosclerosis in the placebo group (decrease in mean luminal diameter: 0.14±0.21 mm for the B1B1 genotype, 0.10±0.20 mm for the B1B2 genotype, and 0.05±0.22 mm for the B2B2 genotype). This association was abolished by pravastatin. Pravastatin therapy slowed the progression of coronary atherosclerosis in B1B1 carriers but not in B2B2 carriers (representing 16 percent of the patients taking pravastatin). There was a significant interaction between pravastatin treatment and decreases in the mean luminal diameter (P = 0.01) and the minimal luminal diameter (P = 0.05). The association of the B1 allele with greater progression of diffuse atherosclerosis (i.e., greater decreases in the mean luminal diameter), as observed in the placebo group, was influenced by the use of pravastatin. In fact, the B1 allele appeared to be associated with less progression in the patients who were receiving pravastatin.

There was a co-dominant relation between the B1 allele and the efficacy of pravastatin in retarding the progression of coronary atherosclerosis. Carriers of two B1 alleles benefited most from treatment with pravastatin: they had significantly less progression of coronary atherosclerosis, as evidenced by smaller decreases in both the mean luminal diameter (P = 0.001) and the minimal luminal diameter (P = 0.002), than their B1B1 counterparts in the placebo group. Furthermore, carriers of only one B1 allele (B1B2) who were receiving pravastatin had significantly less focal atherosclerosis (P = 0.01) than their counterparts in the placebo group. Finally, B2B2 homozygotes had a nonsignificantly greater progression at the end of the study than their counterparts in the placebo group.

Both the association of the CETP TaqIB genotype with the decrease in either the mean luminal diameter or the minimal luminal diameter in the placebo group and the interaction between the genotype and pravastatin treatment remained significant after adjustments were made for the mean luminal diameter (or minimal luminal diameter) at base line, the base-line HDL cholesterol concentration, changes in HDL cholesterol concentrations, and activities of both hepatic lipase and lipoprotein lipase. The precise molecular mechanism that underlies the relation between the CETP gene variant and the angiographic response to pravastatin treatment cannot be deduced from this study. However, it may be related to plasma concentrations of CETP.

The observations suggest that high CETP concentrations, and therefore high levels of CETP activity, result in an enhanced transfer of cholesteryl esters to atherogenic lipoproteins and have negative effects on the structure and function of

the HDL pool, which increases the risk of coronary artery disease. This inference is in agreement with the observation that the pravastatin-induced reduction in CETP concentrations was associated with beneficial angiographic effects in patients who had high CETP concentrations -- that is, those who were homozygous for the B1 allele. In contrast, the reduction in CETP concentrations induced by pravastatin in patients with genetically determined low plasma concentrations of CETP -- that is, those who were homozygous for the B2 allele -- was associated with a lack of retardation of the progression of coronary atherosclerosis. On the basis of these results and the finding of an increased risk of coronary artery disease in subjects who are heterozygous for CETP deficiency, it is believed that a critical concentration of CETP is required for normal reverse cholesterol transport. In contrast, high plasma concentrations of CETP, as seen in placebo-treated B1B1 patients, may promote atherosclerosis by increasing the cholesterol component of atherogenic lipoproteins.

One skilled in the art can apply the knowledge of the CETP allelic differences by applying the techniques as described in the detailed summary section. In this way, one could identify the known allelic differences as described above to identify other allelic differences within the CETP gene. One would then be able to utilize molecular biological techniques to provide a diagnostic test to identify the genotypic differences within a selected group of volunteers or patients. In this way, using the methods for designing and implementing a clinical study described in the detailed description, one could implement a clinical trial to further test the significance of allelic variances on the response to pravastatin, other statins, other cholesterol lowering drugs or other candidates drugs that are known to interact with or affect the CETP gene pathway.

**Example 31**

Angiotensin converting enzyme (ACE)

The ACE polymorphism provides an example of a variance in the non-coding region of a gene with an impact on drug efficacy.

Angiotensin-converting enzyme (ACE) inhibitors, initially developed as antihypertensives, have been shown to reduce mortality in trials of patients with both symptomatic and asymptomatic left ventricular dysfunction and after acute myocardial infarction. An insertion/deletion polymorphism, consisting of a 287-base pair Alu repeat sequence, in intron 16 of the ACE gene, has been shown to predict approximately half of the variance in serum ACE levels between individuals. Homozygotes for the deletion allele (DD) have serum ACE levels twice as high on average as those homozygous for the insertion allele (II), whereas heterozygous (ID)

have intermediate levels. It has been demonstrated that genotype continues to predict residual ACE activity even after acute ACE inhibition with enalapril (Todd et al. Br J Clin Pharmacol 1995; 39:131-4). In a typical pharmacogenetic phase I design, comparing two groups of homozygotes healthy males, DD (n=12) and II (n=11) after genotyping 200 healthy normotensive men, the effect of enalapril, an ACE inhibitor drug, was significantly greater and lasted longer in the men homozygous for the II ACE genotype (Ueda et al. Circulation 1998; 98:2148-2153 ).

**Example 32**

Glycoprotein Integrin beta-3 subunit and Glycoprotein Integrin alpha-2 subunit (GPIIIa/GPIIb)

Glycoproteins IIIa (GPIIIa) and IIb (GPIIb) form the GPIIIa/GPIIb complex that belongs to a class of multisubunit integrin receptors that bind cell adhesion molecules. These receptors are composed of alpha and beta subunits referred to as GPIIb and GPIIIa, respectively. Together the GPIIIa beta and GPIIb alpha subunits form part of the platelet complex receptor, fibronectin receptor, and vitronectin receptor, and play a role in clotting.

The GPIIIa gene encodes a 788 amino acid polypeptide with a 26 residue signal peptide, a 29 residue transmembrane domain near the carboxyterminus and four cysteine rich domains of 33-38 residues each. (Zymrin et al., *J. Clin. Invest. 81:*1470-1475 (1988)). Two different antigenic forms of GPIIIa, alloantigens PlA1 and PlA2 (Platelet Antigen 1 and 2) have been described and can be distinguished using a monoclonal antibody. The most common form of GPIIIa, PlA1, is carried by 98% of the Caucasian population. The rarer form of GPIIIa, PlA2, carries a point mutation or single nucleotide polymorphism at base 192, changing a codon from CTG to CCG thereby causing a leucine to proline substitution at amino acid position 33 (Newman et al., *J. Clin. Invest. 83:*1778-1781 (1989)).

The GPIIb polypeptide is the larger component of the GPIIIa/GPIIb complex and includes two disulfide-linked subunits of 137 amino acids and 871 amino acids respectively. Two antigenic forms of GPIIb, $Bak^a$ and $Bak^b$, have been described and can be distinguished using specific antisera. The rarer form of GPIIb, $Bak^b$, has been shown to have a T to G point mutation that results in an isolucine to serine substitution at amino acid position 843 (Lyman et al., *Blood 75:*2343-2348 (1990)).

The presence of the C-nucleotide at position 192 of GPIIIa DNA can be readily detected by PCR amplification of a region bracketing position 192, followed by MspI digestion of the amplification products, as the C-substitution at that site creates a new MspI restriction site. Alternatively, the sequence at the variance site can be determined using sequencing of the amplification products to identify the nucleotide at the specified position.

The variant GPIIb forms can be detected using similar techniques as for GPIIIa variants by determining the nucleotide at position 2622 (corresponding to amino acid position 843).

It was found that each of the rarer variant sequences described above for GPIIIa and GPIIb correlated with the development of Alzheimer's disease, both separately and together. The variant GPIIIa and GPIIb alleles were found in Alzheimer's patients with an odds ratio of 1.82 and 1.45 respectively as compared to the wild-type alleles. Further, the two variant alleles were found to occur together in Alzheimer's patients as compared to normal subjects with an odds ratio of 3.74.

GPIIIa and GPIIb thus provide examples of variant sequences which result in amino acid substitutions in encoded polypeptides, where the variant sequences are correlated with the development of a disease or condition.

Similarly, other sequence variances in GPIIIa and GPIIb can be anlyzed. In GPIIIa, these include for example, arg62term, leu117trp, asp119tyr, ser162leu, arg214gln, arg214trp, cys374tyr, tro407ala, arg636cys, and ser752tro. For GPIIb, the additional variance include leu183tro, gly242asp, the289ser, glu324lys, erg327his, gly418asp, arg553trm, ile565thr, gln747trp, and ser870term. The possible correlation of these variances with the development of cardiovascular disease can also be determined as for the previously identified variances.

**Example 33**

β2-adrenergic Receptor Polymorphisms and Affects on Outcomes of Congestive Heart Failure

Several variances have been identifies in the gene encoding β-adrenergic receptor. Some of the variances have been shown to affect receptor physiology, and may account, in part, for interpatient variation in the development, progression, and treatment outcomes of congestive heart failure. In a recent study, β-adrenergic receptor polymorphisms were correlated with clinical course of congestive heart failure (CHF) patients. Three amino acid polymorphisms identified in the β-adrenergic receptor were used to stratify patients diagnosed with CHF: Gln27Glu (Glu27 is associated with reduced receptor down-regulation), Arg16Gly (Gly16 is associated with increased down-regulation), and Thr164Ile (Ile 164 is associated with decreased coupling of the receptor with its GTP-binding protein). The allele frequencies of these polymorphisms were similar in normal and CHF patients, suggesting that the polymorphisms are not important in the etiology of CHF.

In a comparison of patients with the 164Ile variance versus the 164Thr variance (more common allele), the investigators determined that survival was greater for the 164Thr variance groupover the study period; unadjusted relative risk

was 4.81 as compared to 3.69. The follow-up survival for individuals with the 164Ile genotype was 42% as compared with 76% for the 164Thr individuals. Although at the time of enrollment the two groups had similar clinical symptoms and other characteristics, there appeared to have been speedier decline in the patients with the Ile164 genotype. An analysis of the other two polymorphic sites (positions 16 and 27), revealed no detectable difference. These data taken together suggest that certain polymorphisms in pharmacologically and/or physiologically relevant proteins may influence the course of disease progression, and establishes the importance of determining genotypic differences to be able to identify individuals with specific genotypes in which earlier aggressive therapy would be warranted.

**Example 34**

Anthracycline Antibiotics

I. *Description of Anthracycline Antibiotics*

The anthracyclines are among the most important cancer drugs due to their broad effectiveness against various carcinomas, sarcomas, leukemias and lymphomas.The anthracycline antibiotics, daunorubicin and doxorubicin, were initially isolated from *Streptomyces peucetious*and have been in clinical use for decades. As a result of the effectiveness of these compounds hundreds of analogs have been produced synthetically or isolated from various microorganisms, including the recently developed compound idarubicin. Other recently isolated anthracyclines include DA-125, moflomycin, SM-5887, IT-62-B, WP631, KRN8602, AD198 and MX2 (all of which show antitumor activity), as well as 3'-O-demethyl mutactimycin, 4-O,3'-O-didemethyl mutactimycin and nothramicin (isolated from non-Streptomyces species). Many other compounds are known to those skilled in the art. These compounds can intercalate into DNA, and interfere with DNA replication and RNA transcription by steric action and by interfering with topoisomerase II function. The anthracyclines are associated with single- and double-stranded DNA strand scission, as well as production of radicals including superoxide, which induce damage to cellular components, including indirect DNA and protein alkylation. Free radical generation is dependent on cellular cytochrome p450 and quinone or hydroxyquinone moieties on adjacent rings in the anthracene backbone. Anthracyclines also bind membranes and alter membrane fluidity and transport, perhaps by radical formation. The anthracyclines are metabolized via hepatic oxidation to polyalcohols , with subsequent deglycosylation, formation of glucuronides and excretion into both bile and urine.

II. *Current Indications for Anthracycline Antibiotics and Derivatives*

Anthracycline antibiotics and derivatives are currently used to treat a broad spectrum of neoplastic diseases including, leukemias, lymphomas, sarcomas, neuroblastomas and cancers of the breast, thyroid, lung, stomach, and urogenital tract (endometrium, ovary, testicle). The synthetic, less cardiotoxic anthracyline derivative, mitoxantrone is indicated for acute nonlymphocytic leukemia (ANLL), and is also active against non-Hodgkins lymphomas and breast cancer.

III. *Limitations of Current Therapies Utilizing Anthracycline Derivative Antibiotics*

The clinical use of anthracycline derivatives is circumscribed by dose-limiting neutropenia and mucositis, and by cardiac toxicity, including an acute syndrome characterized by conduction and rhythm abnormalities or pump failure, and a chronic syndrome of cardiomyopathythat can lead to congestive heart failure. Anthracyclines are administered intravenously on various schedules. In the past dosing was by iv bolus every 3 or 4 weeks, but it has come to be appreciated that repeated small doses or continuous iv infusion is safer, especially in terms of cardiac toxicity, with no evident loss of efficacy. A major limitation of this family of compounds is that a cumulative dose in excess of 550 mg/square meter puts patients at risk of cardiomyopathy and resulting congestive heart failure. In the range of 1 to 10% of patients receiving a cumulative dose of at least 550 mg/square meter develop cardiomyopathy. Cariomyopathy develops in a smaller fraction of patients receiving lower cumulative doses.All clinically tested anthracyclines are effective against some lymphomas and leukemias. Doxorubicin is also effective against certain solid tumors, such as those of the breast and lung, and a wide range of sarcomas. Doxorubicin is the drug of choice for the treatment of metastatic thyroid tumors. It is known to produce severe local toxicity to previously irradiated tissues, even when the two therapies are not administered contemporaneously. Although mitoxantrone treatment produces less nausea, vomiting, and alopecia than doxorubicin, acute myelosuppression and mucositis are frequently observed.

IV. *Impact of Genotyping on Drug Development for Anthracyclines*

The effectiveness of the anthracycline class of chemotherapeutics is believed to be related to its ability to cause DNA damage, either by direct free-radical damage or through the disruption of topoisomerase II function. Other effects of free radicals, which attack a wide range of important biological targets, are also likely to be important. Resistance to treatment can occur through several mechanisms, some of which are well studied. For example, (1) decreased levels of Topoisomerase II are frequently observed in anthracyline resistant cells. Levels of Topoisomerase II

(including TOP2 alpha and TOP2 beta genes) could be influenced by sequence variation or (in cancer tissue) by loss of heterozygosity , affecting interpatient variation in response or toxicity. (2) Topoisomerase III and Topoisomerase III beta levels or function may modulate response to anthracyclines. Anthracyline resistance in experimental systems is often mediated by drug efflux proteins, including the multidrug resistance transporter MDR1 and the multidrug resistance associated protein 1, as well a possibly other members of the ATP-binding cassette family (MRPs 1 through 6). (3) Variation in levels or function of phase I oxidative metabolism, glutathione-S-transferase and peroxidase, , lung resistance related protein (LRP), breast cancer resistance protein (BCRP), and topoisomerase II. As anthracyclin action is exerted through DNA damage, enzymes involved in the detection and repair of DNA damage (such as members of the xeroderma pigmentosum complementation groups (XP), the excision repair cross-complementation groups (ERCC), p53, the ataxia-telangiectasia pathway could also affect efficacy and toxicity. Polymorphisms in any of these gene pathways that affect the enzymatic activity of a gene product, the amount of a gene product, or the interaction of a gene product with anthracycline derivatives would be expected to affect either the initial response to treatment or systemic toxicity.

There is also evidence that anthracyclines are probably less effective in MSI tumors; resistance attributable to impaired ability to detect DNA damage and thence activate apoptosis, and to increased mutation rate.

Impairment of essential free fatty acid metabolism is believed to play a role in therapeutic effect, as well as cardiac toxicity since administration of L-carnitine has been shown to partially reverse cardiac toxicity. The levels of iron, which serves as a mediator of free-radical damage, are also an important factor in cardiac toxicity, since treatment with the iron chelator , ADR-529 is protective. The levels of enzymes controlling oxidative stress, such as superoxide dismutase, are also known to be important determinants of anthracycline toxicity. Doxorubicin and its metabolite doxorubicinol are known to inhibit the action of ion pumps known to be involved in cardiac muscle contraction such as the sarcoplasmic reticulum calcium-dependent ATPase, SERCA1.

Polymorphisms gene products involved in fatty acid metabolism, iron metabolism, calcium concentration, and free radical quenching that alter total enzymatic activity would all be expected to be predictive of toxicity, particularly if the polymorphism is in a gene product whose expression is restricted to or enriched in cardiac tissue (i.e. SERCA1). As an extension, any polymorphism correlated with reduced cardiac function, either manifest or occult, might predispose patients receiving anthracycline antibiotics or derivatives to cardiac toxicity.

**Example 35**

Antimicrotubule Agents

I. *Description of Vinca and Taxus Alkaloids and Derivatives*

The vinca alkaloids, originally extracted from the periwinkle, *Vinca rosea*, and the taxus alkaloid, taxol, isolated from the Western yew, *Taxus brevifolia*, exert their pharmaceutical effects by promoting the destabilization or polymerization, respectively, of microtubule structures involved in cell architecture and division. The vinca alkaloids, and vinorelbine, a newer, better tolerated derivative, share a heterodimeric, heterocyclic structure and bind tubulin with a 1:1 stoichiometry. Binding prevents mitotic spindle function and normal chromosome segregation, leading to apoptotic cell death. Colchicine, an alkaloid extracted from the autumn crocus, *Colchicum autumnale*, shares this mechanism of action, but its use is restricted to the treatment of gout. Taxol, and its more potent derivative, docetaxel, are complex terpenoid compounds that contains a taxane ring nucleus. Treatment with taxus alkaloids causes the accumulation of microtubule aggregates, leading to abnormal cell morphology and arrest of cell division during mitosis. Prior treatment with doxorubicin, which antagonizes cell cycle progression, can reduce therapeutic benefit and increase toxicity. Discodermolide, a polyhydroxylated alkatetraene lactone, binds tubulin at the same location as taxol, causing tubule aggregation in an analogous manner. Epothilone A and B, isolated from the myxobacterium Sorangium cellulosum, and desoxyepothilone B, a less toxic derivative, are also known to exert their antiproliferative effects through microtubule stabilization. Rhizoxin, combretastatin A4, and amphethinile are other recently identified natural product microtubule inhibitors.

II. *Current Indications for Antimicrotubule Agents*

The vinca alkaloids differ significantly in their antitumor effects as well as actions on normal tissues. Vincristine is a standard component of regimens for treating pediatric leukemias and solid tumors and is frequently used to treat adult lymphomas. Vinblastine is utilized primarily for the treatment of lymphomas, neuroblastoma, breast and choriocarcinomas, and as a second-line therapy for various solid tumors. The most important use of vinblastine is conjunction with bleomycin and cisplatin in the curative therapyt of testicular cancer. Vinorelbine has been successfully used as a monotherapy to treat non-small cell lung cancer and breast cancer. Treatment is via weekly intravenous infusion until dose-limiting toxicity is observed or triweekly during vinblastine treatment of testicular cancer.

III. *Limitations of Current Therapies Utilizing Antimicrotubule Agents*

Vinblastine and vinrelbine cause leukopenia and vincristine can cause hypertension through inappropriate vasopressin secretion. Alopecia occurs in approximately 20% of patients receiving vincristine, but is reversible, often without cessation of treatment. All three vinca alkaloids can cause neurotoxicity, but vincristine has predictable cumulative effects. Neurotoxic symptoms include numbness and tingling of the extremities, loss of deep tendon reflexes, and muscle weakness, the latter prompting suspension or reduction of dosing.

Neutropenia and mucositis are frequently observed during taxol treatment, with peripheral stocking-glove sensory neuropathy seen as the dose-limiting toxicity. Many patients experience myalgia for several days after dosing. Dosing can be via short (1-6 hour) or long (72-96 hour) infusions. Pretreatment with corticosteroids or antihistamines has been used to avert hypersensitivity reactions seen with shorter dosing schedules and mucositis is a frequent complication of longer schedules.

Paclitaxel and docetaxel undergo oxidative hepatic metabolism via CYP2C8 and 3A4 and are particularly toxic in patients with reduced liver function.

IV. *Impact of Genotyping on Drug Development for Antimicrotubule Agents*

The effectiveness of antitubule agents is related to their ability to prevent mitosis by affecting spindle assembly and disassembly, by preventing secretory vesicle translocation, and by perturbing normal cellular architecture . Resistance to treatment can occur through alterations in microtubule-associated protein 4 (MAP4), beta tubulin (TUBB), multidrug resistance transporter (MDR1), Bcl-X/Bcl-2 binding protein (BAD), and tyrosine kinase type receptor HER2/NEU. Polymorphisms in any of these gene pathways that affect the enzymatic activity of gene product, amount of gene product, or interaction between gene product and antimicrotubule agent would be expected to affect initial response to treatment.

Since the vinca and taxol classes of antimicrotubule agents have opposing effects on microtubule polymerization state, resistance to one class of agents is often associated with collateral sensitization to the other. Analogously, tubulin or MAP4 polymorphisms that stabilize microtubules would be expected to respond better to taxol therapy than to vinca alkaloid therapy. Microtubules are composed of alpha and beta tubulin subunits and each is encoded by a 15-20 member, dispersed, pseudogene-containing, multigene family restricted in expression to a subset of tissues. For instance, alpha-1 and beta-2 tubulins are restricted to the testis. Polymorphisms in these subunits would be expected to affect primarily the efficacy of antimicrotubule agents for the treatment of testicular cancers.

Taxol derivatives are metabolized primarily by cytochrome P450s CYP2C8 and 3A4. Polymorphisms that affect the enzymatic activity or amount of these gene products would be expected to be predictive of toxicity, especially hepatic and neural. Alpha-3, beta-4, and beta-5 tubulin subunits are restricted to differentiated neural tissues and polymorphisms in these genes affecting protein levels or microtubule agent binding might be predictive of neural toxicity.

**Example 36**

Topoisomerase Inhibitors

I. *Description of Topoisomerase Inhibitors*

Etoposide and teniposide are two semisynthetic glycoside derivatives of podophyllotoxin, a toxic alkaloid from the mayapple, *Podophyllum peltatum*. These compounds have a similar spectrum of antitumor activity and exert their cytotoxic effects by their interaction with cellular topoisomerase II, an enzyme required during DNA replication. The complex between DNA topoisomerase II and etoposide or teniposide is capable of double stranded DNA strand scission, but not strand exchange or ligation. The resulting DNA damage initiates apoptotic cell death. The bulk of administered etoposide is excreted via the kidney unchanged whereas approximately 80% of teniposide is recovered from urine as metabolites. Amsacrine, 4-(9-acridinylamino)-N-(methanesulfonyl)-m-anisidine, is an additional inhibitor of DNA topoisomerase II in clinical trials as part of multiagent induction chemotherapeutic regimen for acute myelogenic leukemias.

Topotecan and irinotecan, derivatives of camptothecin, originally isolated from the bark of *Camptotheca acuminata*, bind DNA topoisomerase I and cause DNA fragmentation and apototic cell death in a manner entirely analogous to the podophyllotoxins. Topotecan is also oxidized by liver cytochromes prior to being excreted via the kidneys. Irinotecan is a prodrug and requires activation by a carboxylesterase to its active metabolite, SN-38. Elimination after hepatic oxidation is via biliary excretion.

II. *Current Indications for Topoisomerase Inhibitors*

Etoposide and teniposide are active against a broad spectrum of tumor types including testicular, small cell lung, various lymphomas, acute granulocytic leukemia, and Kaposi's sarcoma. When combined with cisplatin and bleomycin for testicular tumors and with cisplatin for small cell lung carcinomas, these compounds become the treatment of choice. Administration can be achieved orally (etoposide), but the preferred route is intravenous, with dosing repeated at 2 to 3 day intervals.

Teniposide is better tolerated in patients with compromised renal function than is etoposide.

Topotecan is used primarily for the treatment of ovarian and small cell lung carcinomas and is administered via daily intravenous infusion for 3 or more days. Irinotecan is indicated for the treatment of colorectal cancer and is administered by slow intravenous infusion once weekly for 4 weeks, followed by a two week recovery period. This dosing cycle is repeated until the desired therapeutic endpoint is reached.

III. *Limitations of Current Therapies Utilizing Topoisomerase Inhibitors*

The dose-limiting toxicity for etoposide is leukopenia which peaks approximately two weeks after the onset of treatment. Nausea, vomiting, stomatitis, and diarrhea occur in about 15% of patients receiving etoposide intravenously and 55% who receive it orally. Hepatic toxicity, phlebitis, dermatitis, and reversible alopecia are also observed. Etoposide treatment of childhood acute lymphoblastic leukemia (ALL) has been linked with a secondary myeloid/lymphoid or mixed-lineage leukemia involving a translocation event at 11q23, a region involved in pluripotent stem cell differentiation, that appears 1 to 3 years subsequent to therapy. Teniposide is also used for the treatment of refractory ALL and is associated with myelosuppression, nausea, and vomiting. Undesired effects of topotecan and irinotecan treatment are similar to those of topoisomerase II inhibitors. Nausea associated with irinotecan is often severe enough to require treatment.

IV. *Impact of Genotyping on Drug Development for Topoisomerase Inhibitors*

Resistance to topoisomerase I and II inhibitor therapy can be caused by alterations in topoisomerase activity, topoisomerase levels, or in inhibitor accumulation (MDR1). Multiple amino acid polymorphisms have been reported for both topoisomerases (Table 3) that could potentially affect enzyme activity, drug binding, or protein levels. It has also been shown that cell lines lacking functional Ku86, a protein involved in double-stranded DNA break repair, are hypersensitive to etoposide, suggesting that polymorphisms affecting the levels or activity of this protein in normal tissues or tumor might also be an important determinants of toxicity and efficacy, respectively.

A serious undesired outcome of etoposide treatment of ALL is the development of therapy-related, secondary nonlymphocytic leukemia. Studies have shown that even though etoposide is a known mutagen, mutation rates are not significantly enhanced during treatment, suggesting that patients that acquire secondary leukemia have a natural predisposition . The appearance of secondary

acute myelomonocytic or promyelocytic leukemia is related to DNA translocations involving the 11q23 region that contains the *Drosophila* trithorax homeobox transcription factor homolog MLL or the15q22 region that contains the PML gene, whose product contains transcription factor consensus elements. The latter translocation produces a hybrid protein containing the DNA binding portion of the PML protein fused to the hormone binding portion of the retinoic acid receptor alpha protein. The 11q23 region is known to the contain folate-sensitive fragile site FRA11B, suggesting that polymorphisms in genes involved in folate metabolism may play a contributing role to the appearance of secondary leukemia. Polymorphisms in genes involved in non-pathological DNA rearrangements such as immunoglobulin and T-cell receptor rearrangements (i.e. ataxia telangiectasia, DNA ligase, Ku86, Ku70, etc.), that alter the amount or activity of their gene product, represent candidate genes for association with susceptibility to etoposide treatment-related, secondary leukemia.

**Example 37**

Platinum Coordination Complexes

I. *Description of Platinum Coordination Complexes*

The cytotoxic nature of platinum compounds was first observed in *E. coli* in 1965 and traced to inorganic platinum adducts with ammonium and chloride ions. Of the thousands of platinum derivatives that have been synthesized and tested, cisplatin (cis-diaminodichloroplatinum II) and carboplatin (diaminocyclobutanediacarboxylatoplatinum II), have proven most valuable in the clinic. These platinum complexes seem to enter cells by diffusion and are activated by hydrolysis to a hydrated, cationic diaminoplatinum II species believed to react with nucleic acids and proteins. The N7 position of guanine is particularly prone to modification and intrastrand and interstrand DNA cross-links between proximal guanine bases and adenine and guanine bases are formed. These adducts inhibit DNA replication and transcription, leading to apoptotic cell death. The bulk of compound is excreted unchanged in the urine.

II. *Current Indications for Platinum Coordination Complexes*

Cisplatin produces good response in a broad range of cancers including those of the bladder, head and neck, endometrium, and small cell lung carcinoma. It has been used alone, or in combination with paclitaxel , cyclophosphamide, or doxorubicin for the treatment of ovarian cancers and is curative in combination with bleomycin, etoposide, and vincristine in about 85% of testicular cancers. Platinum compounds are also used as sensitizers for radiation therapy. Administration is via

intravenous infusion subsequent to hydration with saline to minimize renal toxicity. Dosing regimens are 20 mg/m$^2$ for five consecutive days or 100 mg/m$^2$ given once every four weeks.

III. *Limitations of Current Therapies Utilizing Platinum Coordination Complexes*

Marked nausea and vomiting occur in almost all treated patients, but can be controlled with ondasteron or corticosteroids. Renal electrolyte wasting is a frequent occurrence and may lead to tetany and/or seizures. As a result, it is recommended that plasma magnesium levels in patients receiving platinum compounds are routinely monitored. The most serious adverse effects of platinum coordination compound therapy are neurological. Tinnitus and hearing loss in the high frequency range become more frequent and severe with repeated doses and tend to be more pronounced in children. Cisplatin induced neuropathies were first recognized in the early 1980s. Early indications are reduced decreased vibratory sensibility in toes, loss of ankle jerks, and loss of sural nerve response. Peripheral nerves may show axonal degeneration and secondary myelin breakdown. Cisplatin-induced peripheral neuropathy may worsen after discontinuation of treatment, can linger for months to years after cessation of treatment, and can result in death.

IV. *Impact of Genotyping on Drug Development for Platinum Coordination Complexes*

Resistance to platinum compound therapy is generally acquired through the selection of mutant forms of the tumor suppressor protein p53. This protein is involved in the detection of DNA damage and DNA damage-related cell-cycle arrest and apoptosis. Tissue culture studies have shown that these mutants appear to arise spontaneously and become enriched only after platinum treatment. Numerous polymorphisms in the p53 gene have been reported and any that reduce protein amounts or DNA binding activity would be expected to correlate with lower treatment efficacy. As the cytotoxicity of platinum coordination complexes are directly related to their charge and ability to alkylate DNA, enzymes involved in the detection and repair of DNA damage (such as members of the xeroderma pigmentosum complementation groups (XP), the excision repair cross-complementation groups (ERCC), Ku86/70, etc.) could also affect efficacy and toxicity. Elevated levels of some of these enzymes has been found in platinum-resistant ovarian tumor samples. Polymorphisms in any of these gene pathways that affect the enzymatic activity of gene product, the amount of gene product, or the interaction of gene product with platinum-induced DNA adducts would be expected to affect either the initial response to treatment or systemic toxicity.

Neural platinum toxicity appears to be mediated by free, inorganic platinum. Platinum accumulation is greatest in dorsal root ganglia and lowest in neural tissues protected by the blood-brain barrier, consistent with primarily peripheral toxicity. Several agents including glutathione, metallothionein, nerve growth factor, neurotrophin 3, glutamate, and S-2-(3-aminopropylamino)-ethylphosporothioic acid (WR 2721) have shown promise in animal models as a neuroprotectants.

Glutathione is synthesized from the amino acids cysteine, glutamate, and glycine by the consecutive action of gamma-glutamylcysteine synthetase and glutathione synthetase, encoded by single-copy genes and expressed ubiquitously. Polymorphisms in the genes required for glutathione synthesis would be expected to affect primarily the efficacy of platinum compounds. In contrast, metallothioneins are encoded by a 10 to 12 member multigene family. Metallothionein 3 expression is restricted to neural tissue and polymorphisms could be associated with neural toxicity. Polymorphisms in these genes influencing protein levels or activity would be expected to be important predictors of neural toxicity.

**Example 38**

Steroid Hormone Derivatives

I. *Description of Steroid Hormone Derivatives and Related Agents*

Steroidal agents include adrenocorticosteroids and analogs, agents such as aminoglutethimide that regulate the levels of adrenocorticotropic hormone (ACTH), antiestrogens such as tamoxifen, progestins such as hydroxyprogesterone caproate and megestrol acetate, antiandrogens such as flutamide, and gonadotropin releasing hormone (GNRH) and analogs such as goserelin and leuprolide, that decrease secretion of leutenizing hormone (LH) and follicle stimulating hormone (FSH) by the pituitary after long term administration. Depression of FSH and LH levels, in turn, decreases circulating levels of testosterone to castration levels in men and estrogen levels in women to postmenopausal.

Flutamide, tamoxifen, panomifene, and raloxifene are recently developed androgen and estrogen receptor modulators that block the activation of transcription required for the maintenance and function of hormone-responsive tissues. In the absence of androgen- or estrogen-stimulated transcription, proliferation of metastatic prostate and breast cancers is greatly reduced. These agents are usually used in conjunction with cytotoxic chemotherapeutic agents such as alkylating agents, platinum compounds, anthracylines, topoisomerase inhibitors, and microtubule polymerization/depolymerization modulators. Tamoxifen, cyclosporin A, and

verapamil, have all received great clinical attention due to their ability to reverse MDR1-associated drug resistance.

II. *Current Indications Steroid Hormone Derivatives and Related Agents*

Steroidal agents and agents that indirectly affect steroid levels are used against a rather limited number of neoplastic diseases. Corticosteroids such as dexamethasone and prednisone are used alone or in combination with vincristine and anthracylines, with or without methotrexate and asparginase, for the treatment of acute and undifferentiated lymphoblastic leukemia, due to their ability to block lymphocyte proliferation. Aminoglutathimide treatment of metastatic breast cancer with concomitant hydrocortisone supplementation has been largely supplanted by tamoxifen, which acts directly to limit estrogen receptor signaling. The orally available aromatase inhibitors vorozole, letrozole, exemestane, formestane, and anastrozole are currently in development as second-line therapies for the treatment of advanced breast cancer.

Tumors stemming from endocrine tissues and steroid-responsive tissues frequently retain steroid hormone responsiveness initially. This is true for tumors of breast, prostate, testicular, ovarian, and endometrial origin, as well as other less frequent cancers.

Localized prostate cancer is often curable with surgery and/or radiation therapy, but androgen-deprivation therapy becomes the primary hormonal treatment for metastatic disease. Treatment leads to a reduction in symptoms, but is considered palliative since tumors eventually become insensitive. Reduction in serum androgen can be achieved by bilateral orchiectomy, generally reserved for older patients, GNRH analog treatment, and flutamide treatment alone or in conjunction with GNRH analogs. Flutamide decreases the original flare of prostate tumor growth as a result of transient LH increase in GNRH analog monotherapy. Leuprolide and goserelin are administered via intramuscular and subcutaneous injection and are released slowly into the bloodstream; both agents are also indicated for the treatment of breast and endometrial cancers. Flutamide is administered orally, generally three times daily and is currently approved only for use in combination therapy.

Tamoxifen has replaced diethylstilbesterol as the hormonal treatment of choice for estrogen receptor-bearing breast cancers. Both tamoxifen and raloxifene have found recent application for the prevention of postmenopausal decreases in bone density. Tamoxifen is administered by mouth twice daily and is often used for prolonged periods in the context of adjuvant therapy following the initial treatment

of primary breast cancers. The drug is metabolized by oxidation and formation of glucuronides and excreted into the stool via bile.

III. *Limitations of Current Therapies Utilizing Steroid Hormone Derivatives and Related Agents*

As a class, the hormonal agents are extremely well tolerated. Leuprolide, goserelin and flutamide treatment can produce some of the symptoms of menopause including hot flashes, as well as a loss of libido and impotence, but none of these complications is dose limiting. Doses of tamoxifen 20-times the recommended dose are associated with retinal degeneration, but standard doses produce symptoms similar to menopause, weight gain, and gastrointestinal disturbances, none of which is dose limiting. The aromatase inhibitors produce similar side effects. Prolonged use of tamoxifen, such as during adjunct treatment, chemoprevention, or for prevention of postmenopausal osteoporosis, has been linked to the development of endometrial cancers. Patients receiving the standard dose of tamoxifen for two years are twice as likely to develop endometrial cancer than untreated controls.

IV. *Impact of Genotyping on Drug Development for Steroid Hormone Derivatives*

Antiestrogen therapy in the context of chemotherapy is generally indicated only for estrogen-receptor bearing tumors. Estrogen receptor polymorphisms that affect protein levels, DNA or estrogen binding, or interaction with other transcription factors would be expected to correlate with treatment outcome. More-specifically, decreases in any of these parameters should decrease efficacy. Expression of epidermal growth factor receptor (EGFR) and tyrosine kinase-type cell surface receptor HER2/NEU correlates with poor response to tamoxifen even in estrogen receptor positive tumors, but neither EGFR or HER2/NEU appear to be amplified during the course of treatment. As noted above, HER2/NEU expression also correlates with poor prognosis during treatment with antimicrotubule agents, suggesting that ectopic or enhanced expression of growth factor receptors can overcome the growth inhibition caused by cytotoxic agents.

Steroid hormone derivatives are metabolizes via cytochrome P450 and flavin containing monooxygenases and by conjugation to sulfates and glucuronates for elimination. Oxidative metabolism of tamoxifen by liver microsomal fractions has been well characterized and involves the formation of 4-hydroxyl, 4'-hydroxyl, N-oxide, N-desmethyl, 3,4-dihydroxyl, and 3,4-epoxyl derivatives. The latter, reactive epoxide species is formed in large amounts in rats, but not mice or humans, and is thought to account for increased liver carcinogenesis in this species. Formation of the N-oxide is believed to be mediated by a flavin containing monooxygenase

(FMO), while other reactions appear to be carried out by cytochromes, especially CYP3A4 (N-demethylation) and 2D6 (4-hydroxylation). Polymorphisms in genes encoding FMO enzymes 1 and 3-5 (FMO2 is inactive), or CYP3A4 and 2D6, or sulfotransferases, or glycosyltransferases that affect protein amount or activity would be expected to influence efficacy by increasing or decreasing elimination.

Polymorphisms in the gene(s) encoding the enzyme(s) responsible for 3,4-epoxide formation that affect protein amount or activity would be expected to correlate with the mutagenic effects of tamoxifen, especially with the occurrence of treatment-related endometrial cancers.

**Example 39**

Inhibitors of Signal Transduction

I. *Description of Signal Transduction Inhibitors*

Signal transduction is the processes whereby external cellular stimuli are converted into changes in protein expression. The usual chain of events is (1) interaction of a ligand with a cell surface receptor, (2) ligand-induced changes in the three-dimensional structure of the receptor, including dimerization, that are transmitted to the cytoplasmic face of the plasma membrane, (3) transmission of these changes to transcription factors in a complex, multienzyme cascade that generally involves a change in the phosphorylation state of enzymes in the cascade, modulating their activity or ability to interact with other proteins in the cascade, and (4) modulation of the availability or activity of gene specific transcription factors through changes in phosphorylation status, oligomeric state, cellular localization, synthesis, or degradation.

Changes in the signal transduction process play a pivotal role in the etiology of neoplastic transformation and disease. The enzymes involved are usually members of a closely-related, multienzyme family, that display complex temporal regulation during development and are differentially expressed in normal tissues. Recent advances in understanding the molecular biology of these pathways and technical breakthroughs in both combinatorial chemistry and high-throughput screening have added novel, synthetic agents to the small number of known, naturally occurring signal transduction inhibitors.

Current signal transduction inhibitors affect phosphorylation and dephosphorylation steps associated with receptor and soluble kinases as well as protein phosphatases. Many subtype selective and nonselective inhibitors of protein kinase C (PKC) were initially isolated from natural sources. These include the protein kinase C inhibitors staurosporin , herbimycin A, lavendustin A, and erbstatin, originally isolated from various *Stremptomyces* species; the tyrosine

kinase inhibitors emodin, cytovaricin B, angelmicin B, geldanamycin, and genistein isolated from *Talaromyces, Streptomyces,* and *Lupinus* species; the the phosphatidylinositol 3-kinase inhibitor wortmannin isolated from *Talaromyce flavus*; and the protein phosphatase inhibitor okadaic acid isolated from the black sponge, *Prorocentrum oncauum*. Erbstatins, which bind the ATP-binding site of PKC, also have activity against topoisomerases I and II. Numerous synthetic derivatives of these compounds that enhance stability and availibility, as well as novel compounds, have been produced and assayed in biological systems more recently. These include the protein kinase C inhibitors L86-8275, H7, LY333531 (Eli Lilly), safignol (Sphinx/Eli Lilly) and CGP41251(Ciba-Geigy); the tyrosine kinase inhibitors SU5416 (Sugen), specific for VEGF receptor-associated tyrosine kinase; ZM 252868, ZD1839 (Zeneca), PD153035 (Parke-Davis), and CGP 52411 (Ciba-Geigy), specific for EGF receptor-associated tyrosine kinase, CEP-701 (KT-5555) and K252a, specific for TRK-type receptor-associated tyrosine kinase, KN-62, specific for $Ca^{++}$/calmodulin-dependent protein kinase II; tyrosine kinase inhibitors of the tyrphostin class, as exemplified by AG1714 (4-nitrobenzylidene malononitrile ); the phosphatidylinositol 3 kinase inhibitors of the 3-deoxy-D-myoinositol 1-phosphate/1-phosphonate class; the protein serine/threonine phosphatase inhibitor endothall; and the tyrosine phosphatase inhibitor bis(maltolato)oxovanadium(IV).

Bryostatin, a macrolactone originally isolated from marine sponges, inhibits signal transduction through an as yet unknown mechanism, but likely to involve PKC isozymes.

The growth factor 1 family of transmembrane receptors, including epidermal growth factor receptor (EGFR) and members of the ERB-B family, are overexpressed in a wide variety of solid tumors, particularily squamous cell carcinomas of the head and neck, lung, and cervix. The oncogene CBL, implicated in pre/pro B-cell lymphomas, and its cellular counterpart C-CBL, mediate the recycling/degradation of EGFR members. EGFR members are already important therapeutic targets and C-CBL represents an important future drug target.

Mutations in members of the RAS G protein superfamily are the most common initiators of neoplastic transformation, occurring in approximately 40% of colorectal cancers, 90% of pancreatic cancers, 30% of lung adenocarcinomas, and 25% of acute myeloid leukemias. Inhibitors of RAS farnesylation are currently in clinical trials.

RAS control of cell cycle decisions and of transcription factor phosphorylation by Janus kinase (JUNK) is mediated by at least four distinct signaling pathways including the mitogen-activated protein kinases (MAPKs) and

phosphatidyl inositol 3 kinases (PIK3s). Inhibitors designed to members of these multigene families are in early development, show great promise. Silymarin, a flavonoid antioxidant isolated from milk thistle, is a MAPK modulator and has shown great efficacy in the chemoprevention of skin cancer in a mouse model.

II. *Current Indications for Signal Transduction Inhibitors*

Many of the signal transduction inhibitors listed above show promising antineoplastic activity in tissue culture and animal models, but only a few compound are currently in early clinical development. The staurosporine derivative UCN-01 is being tested for activity in advanced or refractory solid tumors, lymphoproliferative disorders, and lymphoid malignancies and SU5416 is being tested in combination therapy with 5-fluorouracil and leucovorin for metastatic colorectal cancer. Bryostatin is in early clinical trials alone, or in combination with cisplatin and paclitaxel for refractory and advanced malignancies including unresectable stomach, esophagus, anus, prostate, or non-small cell lung cancer. It is anticipated that upon further development, members of this class of agents will be indicated for the treatment of a broad range of neoplastic diseases. The phosphatidylinositol 3 kinase inhibitor, wortmannin, has been shown to be active as a radiosensitizer *in vitro*, suggesting potential utility in the radiation therapy of tumors.

III. *Impact of Genotyping on Drug Development for Signal Transduction Inhibitors*

Activation of the MAPK pathway has been shown to correlate with poor prognosis for prostate cancers as well as resistance to androgen ablation therapy. In both colorectal and breast tumors, expression in malignant tissue was elevated while expression in surrounding tissues was normal. The oncogenic potential of MAPKs has been demonstrated in a mouse model, where a lysine to glutamate mutation appears to cause cellular transformation. These findings all highlight the important role of the MAPK pathway in the initiation and progression neoplastic disease.

There is also evidence for the direct involvement of the PKC pathway in neoplastic disease: a point mutation at position 294 of alpha-protein kinase C, leading to an aspartic acid to glycine substitution, has been linked to pituitary tumor invasiveness.

Overexpression of EGFR has been strongly associated with the transition from superficial to invasive bladder cancers. Enhanced cellular motility is a prerequisite to invasion and can be inhibited in an *in vitro* model by wortmannin, a specific inhibitor of phosphatidylinositol 3 kinase, implicating this class of enzymes in bladder cancer progression. The structurally related ataxia telangiectasia gene product, when mutated, causes a predisposition to malignancy.

Polymorphisms in genes encoding receptors and proteins of the signal transduction pathway as detailed above and in Tables 1-3, or related proteins yet to be discovered, which influence protein amounts, activity, interaction with other proteins or drugs, would be expected to have prognostic value for risk assessment, treatment efficacy, and toxicity.

**Example 40**

Inhibitors of Cell Cycle Control

a) Description of Cell Cycle Control Inhibitors

The control of cell cycle progression and division is through a complex signaling pathway, such as described above, at the heart of which are the cell division cycle (CDC) proteins, CDC kinases (CDCKs), cyclins, cyclin-dependent kinases (CDKs), and cyclin-dependent kinase inhibitors. All exist as members of multigene families that show temporally regulated and tissue-specific expression. Ubiquitin ligases and the ubiquitinated protein proteolysis pathway are involved in modulating cyclin levels during the cell cycle.

Interest in this class of macromolecules as drug development targets was sparked by the observation that the level of various cyclin-dependent kinases and cyclin-dependent kinase inhibitors differed between normal tissues and a wide variety of tumor types and could be prognostic of treatment outcome. Low to absent levels of the cyclin-dependent kinase inhibitor 1B (p27, KIP1) has been shown to be predictive of unfavorable prognosis in a variety of tumors. Ectopic expression of KIP1in human brain tumor cells has been shown to reverse some of the changes of neoplastic transformation. In contrast, elevated levels of the cyclin-dependent kinase inhibitor 2A (p16, INK4, MTS1), is associated with progression and unfavorable prognosis in prostate and ovarian cancers. This protein is also the target of frequent somatic mutation.

Several natural and synthetic inhibitors of CDK and CDCK function have been isolated. These include flavopiridol, butyrolactone I, and the purine derivatives aminopurvalanol, olomoucine, and roscovitine.

II. *Current Indications for Cell Cycle Control Inhibitors*

The CDK and CDCK inhibitors listed above have demonstrated efficacy in a number of transformed cell types in tissue culture, but only flavopiridol is advanced clinical development for refractory and recurrent colorectal cancer, adenocarcinoma of the prostate, lymphocytic leukemia, and non-Hodkin's and mantle cell

lymphomas, either as a monotherapy or in combination with taxol and cisplatinum compounds. It is anticipated that upon further development, members of this class of agents will be indicated for the treatment of a broad range of neoplastic diseases and hyperproliferative disorders such as psoriasis and restenosis.

III. *Impact of Genotyping on Drug Development for Cell Cycle Control Inhibitors*

Expression of cyclin-dependent kinase inhibitor 1A (p21/WAF1) is induced by the tumor suppressor protein p53 in response to DNA damage, thereby playing a direct role in mediating p53-induced G1 arrest. Two polymorphisms in the p21 gene, a serine to argenine change at codon 31 a C toT transition in non-coding sequence, show increased prevalence in prostate adenocarcinoma and squamous cell carcinoma of the head and neck. Similarly, low to absent levels of the cyclin-dependent kinase inhibitor, p27/KIP1are associated with poor clinical outcome in gastric and colorectal cancers.

Cyclin D1expression levels have also been shown to correlate with progression and prognosis in non-small cell lung cancers, estrogen receptor-positive breast cancers, esophageal cancer, and gastric cancers. However, the correlation can be positive or negative, depending upon cancer type, making it likely that cyclin D1 levels are not the directly responsible for neoplastic transformation in these tumors, and that they are a poor prognostic indicator for tumors in general. However, analysis of patients diagnosed with squamous cell carcinomas showed that G/G homozygotes of the silent G/A polymorphism in exon 4 of cyclin D1 tend to exhibit less differentiated tumors and have shorter remission times than G/A heterozygotes and A/A homozygotes. These findings carried over to various tumor subtypes, including laryngeal and pharyngeal.

Polymorphisms in genes encoding cell cycle checkpoint proteins, and proteins involved in cell cycle progress as detailed above and in Tables 1-3, or related proteins yet to be discovered, which influence protein amounts, activity, interaction with other proteins or drugs, would be expected to have prognostic value for risk assessment, treatment efficacy, and toxicity.

**Example 41**

Angiogenesis Inhibitors

i. Description of Angiogenesis Inhibitors

The utility of angiogenesis inhibitors for the treatment of solid tumors was first recognized by Folkman and colleagues in 1980. Angiogenesis, the creation of vasculature, is a process that insures that tissues and organs are adequately supplied

with oxygen and nutrients and that toxic metabolites are efficiently removed. Angiogenesis involves the release of growth factor gradients by inadequately supplied tissue, response to these factors mediated by receptors in surrounding vasculature, and proteases and adhesion molecules involved in tissue remodelling. Angiogenesis and neovascularization, inappropriate or abnormal angiogenesis, can be induced by a number of pathological conditions, usually in the context of hypoxia or inflammation.

As rapidly growing cell masses, solid tumors require a constant, plentiful supply of oxygen and nutrients. In larger tumors, perfusion is often inadequate, causing hypoxia and central necrosis. Various classes of compounds including inhibitors of signal transduction (i.e. LY333531), inhibitors of growth factor receptors (i.e. SU5416), protease inhibitors (i.e. KB-R7785, marimastat), and adhesion inhibitors (castanospermine) have shown activity in various models of angiogenesis and against multiple solid tumor types. Compounds showing promise in model systems or currently in development include the peptides aplidine, vascular endothelial growth inhibitor (VEGI), brain-specific angiogenesis inhibitor (BAI1), K1-5 (kringles 1-5 of plasminogen), U-995 (shark cartilage derived), endostatin, angiostatin, an antibody against vascular endothelial growth factor, and macrophage inflammatory protein 2 (MIP2, GRO2); the steroids and terpenoids squalamine, vitamin D3, and retinoic acid; the antibiotics clarithromycin and combretastatin A4; and the synthetic compounds SU5416 (Sugen), TNP-470, COL-3, IM862, PTK787/ZK222584 (Zeneca), CT-2584, KB-R7785, LY333531 (Eli Lilly), BPHA (Shionogi), carboxyamidotriazole, 5,6-dimethylxanthenone-4-acetic acid, and alpha-difluoromethylornithine, an inhibitor of polyamine synthesis.

II. *Current Indications for Angiogenesis Inhibitors*

Clinical trials of antiangiogenesis agents are underway for a wide variety of refractory and recurrent solid tumor types including Kaposi's sarcoma, non-Hodgkin's lymphoma, astrocytoma, glioblastoma, oligodendroglioma, ovarian, prostate, and renal tumors.

III. *Impact of Genotyping on Drug Development for Angiogenesis Inhibitors*

Vascular endothelial growth factor (VEGF) gene expression is increased in k-RAS transformed colorectal cells and VEGF expression is required for efficient tumor formation in nude mice but not for cell immortality. VEGF expression is associated with the progression, invasion and metastasis of colorectal cancer and overexpression of VEGF mRNA in the primary tumour is closely correlated with poor prognosis. High pretreatment serum VEGF is associated with poor response to

treatment and unfavorable survival in patients with small cell lung cancer treated with cisplatin and etoposide combination chemotherapy. These findings suggest the importance of this growth factor in tumor proliferation and implicate polymorphisms in VEGF proteins and VEGF receptor as potentially important determinants of prognosis, treatment efficacy, and toxicity.

The plasminogen-derived antiangiogenic peptide, angiostatin binds vitronectin and induces focal adhesion kinase (FAK1) activity in cell culture. FAK1 normally becomes phosphorylated only in response to cell-cell contact or treatment with peptide hormones including cholecystokinin, bombesin, and vasopressin. This observation suggests that the biological effects of angiostatin may relate to subversion of adhesion plaque formation in endothelial cells. The collagen XVIII-derived antiangiogenic peptide, endostatin binds fibulins 1 and 2 and also induces FAK1 activity.

Macrophage metalloproteinase (HME/MMP12) expression levels in hepatocellular tumors correlate well with angiostatin levels, which in turn were inversely correlated with poor survival. Transforming growth factor-beta 1, a key mediator of tumor angiogenesis, inhibits the generation of angiostatin in a pancreatic carcinoma cell line through modulation of the plasminogen activator/plasminogen activator inhibitor system. Generation of angiostatin may also involve an as yet unidentified, secreted disulfide reductase.

Polymorphisms in genes listed above or in Tables 1-3 and including similar genes not yet discovered that encode vascular growth factors, their receptors, and in enzymes involved in their processing that affect enzyme amounts, activity, or interaction with drug molecules could potentially affect neoplastic disease risk and prognosis as well as antiangiogenic treatment efficacy and toxicity.

**Example 42**

Protease Inhibitors

I. *Description of Protease Inhibitors*

Extracellular proteases play a crucial role in normal tissue remodeling during embryogenesis, growth, and wound healing by modulating the maturation and degradation of growth factors and extracellular matrix components such as elastin and collagen. Proteases play a role in angiogenesis—the potent inhibitors angiostatin and endostatin are proteolytic fragments of plaminogen and collagen 18A1, respectively. Extracellular proteases are involved in the progression of multiple pathological conditions such as osteoporosis and multiple inflammatory disorders including rheumatoid arthritis, multiple sclerosis, and nephritis.

Tumor metastasis, the migration of cells from the primary tumor to distal sites via the lymph or blood vessels, is mechanistically similar to the migration of lymphocytes from the lymph nodes to sites of inflammation, a process known to rely on the action of zinc requiring matrix metalloproteases (MMPs) and to be regulated by corresponding tissue inhibitors of metalloproteinases (TIMPs). Both matrix metalloproteases and their inhibitors occur in large, dispersed multigene families. Levels of MMP 1 and TIMP 1 correlate with metastatic potential and poor treatment outcome in breast, gastric, and colorectal cancers; levels of MMP 2 and TIMP 2 correlate with metastatic potential and poor treatment outcome in renal, urothelial, bladder, and colorectal cancers. Invasion of smooth muscle cell layers by tumor cells is inhibited by TIMPs and transfection of human breast cancer cells with TIMP4 reduces their growth and metastatic potential, suggesting direct involvement of metalloproteases in metastasis.

Several matrix metalloprotease have shown promising activity in tissue culture and *in vivo* models of metastasis including biphenyl sulfonyl-phenylalanine hydroxiamic acid (BPHA), KB-R7785, and R-94138; several inhibitors including marimastat, batimastat, and AG3340 (Agouron) are in various stages of clinical development.

II. *Current Indications for Protease Inhibitors*

Clinical trials of protease inhibitors in progress target advanced lung cancers including small cell and non-small cell; supratentorial glioblastoma multiforme; gliosarcoma; gastric, pancreatic, and metastatic breast cancers; and combination therapy with mitoxantrone and prednisone for hormone refractory prostate cancer. Batimastat has shown promise for the treatment of malignant pleural effusion.

Because proteinase inhibitors are not cytotoxic, their use in anticancer therapies has been in combination with cytotoxic agents such as anthracycline antibiotics, microtubule inhibitors, topoisomerase inhibitors, etc., where they inhibit tumor growth indirectly through their antiangiogenic effects and tumor metastasis directly by inhibition of enzymes required for tumor dispersion.

As metalloproteases have been implicated in tumor metastasis, protease inhibitors may find widespread application for the prophylactic treatment of primary tumors during standard chemotherapeutic regimens to prevent the migration of (resistant) tumor cells to secondary sites.

III. *Limitations of Current Therapies Utilizing Protease Inhibitors*

Symptoms reported by patients with various malignancies during trials of marimastat included severe joint and muscle pain which were debilitating in >60%

of patients at doses >50 mg twice daily. These symptoms were reversible on discontinuation of the drug, and their incidence was been decreased by reducing the dose to10 mg twice daily.

IV. *Impact of Genotyping on Drug Development for Protease Inhibitors*

Protease inhibitors have great potential in the treatment of neoplastic disease through their apparent ability to inhibit tumor invasion and dispersion. The protease/protease inhibitor systems that have been implicated in this process include the matrix metalloproteinases (MMPs) and their corresponding tissue inhibitors of metalloproteinase (TIMPs), the cathepsins (CTSs), and plasminogen activator (PLAU), plasminogen activator receptor (PLAUR), and plasminogen activator inhibitor (PAI1). As proteases are also involved in the inhibition of the angiogenesis required for tumor growth by releasing the potent inhibitors of angiogenesis, endostatin and angiostatin from collagen and plasminogen, greater understanding of the protease biology involved in these opposing processes will be required before protease inhibitor therapy can realize its full potential.

Serum levels of PLAU, PAI1, and PLAUR are predictors of progression and prognosis in prostate and gastric cancers: higher levels correlate with poor outcome and prophylactic chemotherapy after resection may be warranted for patients displaying high levels. Similarly, the five year relapse rate of patients having node-negative breast cancer with low PAI1 and low cathepsin D (CSTD) was 13% while patients who had greater than the median value for both of these molecules had a 5 year relapse rate of 40%. These data would indicate that at least two different protease systems are active in spread of node negative breast cancer and that measurement of CSTD and PAI1 levels may aid in the decisions to be made when offering adjuvant treatment to these patients. Cathepsin B (CTSB) is overexpressed in tumors of the lung, prostate, colon, breast, and stomach.
Abundant extracellular expression of CTSB protein was found in 29 of 40 (72.5%) of esophageal adenocarcinoma specimens by use of immunohistochemical analysis.

A single nucleotide insertional polymorphism at -1607 in the promoter of matrix metalloproteinase 1 (MMP1), where an additional guanine (G) creates an Ets transcription factor binding site, creates an allele that displays significantly higher transcription in normal fibroblasts and in melanoma cells. This polymorphism occurs in the normal population with a frequency of 30%. In contrast, in eight tumor cell lines, this frequency increased to 62.5% ($P < 0.0001$), perhaps because increased levels of MMP1 allow more aggressive matrix degradation, thereby facilitating cancer progression.

Polymorphisms in genes listed above or in Tables 1-3 and including similar genes not yet discovered that encode proteases, their substrates (including adhesion proteins), their inhibitors, and in enzymes involved in their processing that affect enzyme amounts, activity, or interaction with drug molecules could potentially affect neoplastic disease risk and prognosis as well as protease inhibitor treatment efficacy and toxicity.

**Example 43**

Use of Genotype Information for the Identification of Candidates for Prophylactic Therapy

i) Occult Disease Detection and Prophylaxis

The early detection and treatment of neoplastic disease greatly improves prognosis—the prognosis for breast cancer chemotherapy is inversely related to lymph node involvement. Genotyping of polymorphisms known to be associated with increased risk for neoplastic disease would warrant careful monitoring or prophylactic treatment. Patients and practitioners must carefully weigh the benefits and associated undesired toxicities of prophylactic treatment against the risk of disease onset and response to conventional therapies.

Great advances in linking genetic polymorphisms to cancer risk have been made in recent years. Most link polymorphisms in genes involved in drug and xenobiotic metabolism (primarily phase I metabolism) to the appearance of various cancers. As environmental risk factors can be controlled, they can be viewed as modulators of genetic polymorphisms involved in innate risk. These include, but are not restricted to, the genes in the table below, which are known to be polymorphic and polymorphisms have been linked to innate or environmentally induced cancer risk in the scientific literature.

| | Associated Polymorphic Gene | | | | |
|---|---|---|---|---|---|
| **Cancer** | **Name** | **GID** | **OMIM_ID** | **VGX_Symbol** | |
| Bladder | GSTM1 | J03819 | 138352 | GEN-9D | 8824515 |
| | NAT2 | D90041 | 243401 | GEN-466 | 10510890 |
| | CYP2D6 | X08007 | 124031 | GEN-1FE | 8824515 |
| Breast | CYP17 | M14564 | 202110 | GEN-2Z | 99415566<br>10519398<br>10404084<br>9950238<br>9067272 |

| | | | | | |
|---|---|---|---|---|---|
| | CYP1A1 | K03191 | 108330 | GEN-9E | 10468307<br>10519398 |
| | COMT | M58525 | 116790 | GEN-3S | 10519398 |
| | NAT2 | D90040 | 243400 | GEN-465 | 10389748 |
| | HRAS | J00277 | 190020 | GEN-MH8 | 8385520 |
| | VDR | J03258 | 601769 | GEN-2J | 10344739<br>9613456 |
| Colorectal | MTR | U73338 | 156570 | GEN-69 | 10498402 |
| | NAT1 | D90042 | 108346 | GEN-465 | 7627961 |
| | APC | M74088 | 175100 | GEN-3MW | 9973276<br>9869603 |
| | GSTM1 | J03818 | 138351 | GEN-9D | 10445390 |
| | HRAS | J00279 | 190022 | GEN-MH10 | 2887194 |
| | MTHFR | U09806 | 236250 | GEN-4FZ | 9067278<br>8895734 |

| | Associated Polymorphic Gene | | | | |
|---|---|---|---|---|---|
| **Cancer** | **Name** | **GID** | **OMIM_ID** | **VGX_Symbol** | |
| Gastric | NAT1 | D90043 | 108347 | GEN-466 | 10585581 |
| | MYCL1 | M19720 | 164850 | GEN-MK0 | 9635822 |
| | MUC6 | U97698 | 158374 | GEN-LTG | 9419405 |
| | MUC1 | X52228 | 158340 | GEN-33N | 9076520 |
| Glioma | RB1 | M33647 | 180200 | GEN-2K1 | 9210953 |
| Lung | CYP1A1 | K03191 | 108330 | GEN-9E | 9610791<br>10506106 |
| | DIA4 | J03934 | 125860 | GEN-12L | 10397241<br>8528266 |
| | MPO | X04876 | 254600 | GEN-PS | 9371491 |
| | GSTM1 | J03817 | 138350 | GEN-9D | 10506106<br>7728947 |
| | MYCL1 | M19721 | 164851 | GEN-MK1 | 1345822 |
| | GSTM3 | J05459 | 138390 | GEN-17O | 7728947 |
| | OGG1 | Y13277 | 601982 | GEN-9O | 9935223 |
| Melanoma | HRAS | J00278 | 190021 | GEN-MH9 | 2572539 |
| Myeloid Leukemia | IFNB | V00546 | 147640 | GEN-TV | 7912973 |
| Oral | GSTP1 | X06547 | 134660 | GEN-19N | 10376763 |
| | CYP2D6 | X08006 | 124030 | GEN-1FE | 9825835 |
| Ovarian | EPHX | L25878 | 132810 | GEN-29Z | 8944076 |
| Prostate | CYP17 | M14564 | 202110 | GEN-2Z | 10469617 |
| | NAT1 | D90041 | 108345 | GEN-464 | 10211944 |
| | VDR | J03259 | 601770 | GEN-2J | 8797574 |
| Testicular | WT1 | X51630 | 194070 | GEN-32A | 8056449 |

Table: Polymorphic Genes Linked to Cancer Risk. Column 1, labelled "Cancer" shows commonly observed neoplastic diseases classified by organ or cell-type. Genes for which polymorphisms are linked to cancers listed in column 1 are under the broad heading "Associated Polymorphic Gene." These genes are identified by systematic name, "Name;" Genebank identifier, "GID;" Online Mendelian Inheritance in Man identifier, "OMIM_ID;" and internal, Variagenics, Inc. identifier, "VGX_Symbol." In addition, the PubMed database identifier, "PMID," allowing identification of pertinent literature is also given. Worldwide web addresses for the databases mentioned are in the "DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS" section under the subheading "Online Databases."

It is likely that the interplay of multiple polymorphic and non-polymorphic genes is involved in the process of neoplastic transformation with both positive and negative risk associations. But, a patient having several predisposing factors for a given cancer type listed in column 1 of the table above will be at greater risk than a patient having fewer. For example, patients with polymorphisms in glutathione-S-transferase (GST) M3 that reduce expressed levels are more likely to develop lung cancer if they also express low levels of GST M1. One skilled in the art will recognize that knowledge of risk ratios associated with various gene polymorphisms and neoplastic diseases will allow medical practitioners to determine whether prophylactic treatment , including change of habits or environment, preventative chemotherapy, careful monitoring for signs of disease, and prophylactic surgery, are warranted and advisable. Multiple risk-associated allelic loci can be genotyped to direct a course of prophylactic treatment in much the same manner as a high cholesterol count in a blood test carries an increased risk of heart disease and may warrant treatment with a statin-type drug (HMGCoA inhibitor).

It will also be recognized by one skilled in the art that factors including age, sex (in the case of non-gender specific cancers), ethnic background, and environment (including diet, smoking, alcohol consumption, and diet) impact risk determinations and great care must be exercised in extrapolating from one population to another.

II. *Post-Treatment Prophylaxis*

Notes: Aim is to forestall onset of new disease after successful initial therapy--correlation of tumor genotype with metastatic potential.

KIP1 polymorphisms, WAF1 polymorphisms, EGFR polymorphisms, ERBB2 polymorphisms

5

Other Embodiments

The invention described herein provides a method for identifying patients with a risk of developing neurological disease or dysfunction by determining the patients allele status for a gene listed in Tables 1-6, 11-17, and 18-23 and providing a forecast of the patients ability to respond to or tolerate a given drug treatment. In particular, the invention provides a method for determining, based on the presence or absence of a polymorphism, a patient's likely response to drug therapies of neurological disease or dysfunction. Given the predictive value of the described polymorphisms a candidate polymorphism is likely to have a similar predictive value for other drugs acting through other pharmacological mechanisms. Thus, the methods of the invention may be used to determine a patient's response to other drugs including, without limitation, antihypertensives, anti-obesity, anti-hyperlipidemic, or anti-proliferative, antioxidants, or enhancers of terminal differentiation.

In addition, while determining the presence or absence of the candidate allele is a clear predictor determining the efficacy of a drug on a given patient, other allelic variants of reduced catalytic activity are envisioned as predicting drug efficacy using the methods described herein. In particular, the methods of the invention may be used to treat patients with any of the possible variances, e.g., as described in Table 3 of Stanton et al., U.S. Application No. 09/300,747.

In addition, while the methods described herein are preferably used for the treatment of human patients, non-human animals (e.g., dogs, cats, sheep, cattle and other bovines, swine, and apes and other non-human primates) may also be treated using the methods of the invention.

It will be readily apparent to one skilled in the art that varying substitutions and modifications may be made to the invention disclosed herein without departing from the scope and spirit of the invention. For example, using other compounds, and/or methods of administration are all within the scope of the present invention. Thus, such additional embodiments are within the scope of the present invention and the following claims.

The invention illustratively described herein suitably may be practiced in the absence of any element or elements, limitation or limitations which is not specifically disclosed herein. Thus, for example, in each instance herein any of the terms "comprising", "consisting essentially of" and "consisting of" may be replaced with either of the other two terms. The terms and expressions which have been employed are used as terms of description and not of limitation, and there is no intention that in the use of such terms and expressions of excluding any equivalents

of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention has been specifically disclosed by preferred embodiments and optional features, modification and variation of the concepts herein disclosed may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention as defined by the appended claims.

In addition, where features or aspects of the invention are described in terms of Markush groups or other grouping of alternatives, those skilled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group or other group.

| Table 1. C. ology Gene List | | | | | | |
|---|---|---|---|---|---|---|
| **Class** | **Pathway** | **Function** | **Name** | **OMIM** | **GID** | **Locus** |
| | **Drug Uptake and Export** *(additional genes in Toxicology)* | **Transporter** | multidrug resistance associated protein MRP1 | 158343 | L05628 | 16p13.1 |
| | | | multidrug resistance associated protein MRP2/CMOAT | 601107 | NM_000392 | 10q24 |
| | | | ATP-binding cassette, sub-family C (CFTR/MRP), member 3/CMOAT2 | ****** | NM_003786 | ****** |
| | | | ATP-binding cassette, sub-family C (CFTR/MRP), member 4/MOATB | ****** | NM_005845 | ****** |
| | | | ATP-binding cassette, sub-family C (CFTR/MRP), member 5/SMRP | ****** | NM_005688 | ****** |
| | | | ATP-binding cassette, sub-family C (CFTR/MRP), member 9/SUR2 | 601439 | NM_005691 | ****** |
| | | | multidrug resistance protein MDR1 | 171050 | X96395 | 7q21.1 |
| | | | multidrug resistance protein MDR3/P-glycoprotein 3/PGY3 | 602347 | X06181 | 7q21.1 |
| | | | Human sorcin/SRI | 182520 | L12387 | 7q21.1 |
| | | | Placenta-specific ATP-binding cassette transporter/ABCP | 603756 | NM_004827 | 4q22 |
| | | | anthracylcline resistance-related protein/ARA | 603234 | NM_001171 | 16p13.1 |
| | | | sulfonylurea receptor (hyperinsulinemia)/SUR | 600509 | NM_000352 | 11p15.1 |
| | | | Solute carrier family 29, member 1/SLC29A1/ENT1 | 602193 | NM_004955 | 6p21.2-p21.1 |
| | | | Solute carrier family 29, member 2/SLC29A2/ENT2 | 602110 | X86681 | 11q13 |
| | | | Glutathione-S-transferase 6 | 138391 | ****** | ****** |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Drug Metabolism** *(additional genes in Toxicology* | **Drug Inactivation** *(additional genes in Toxicology)* | **Glutathione** | Glutathione-S-transferase, alpha 1/GSTA1 | 138359 | L13269 | 6p12.2 |
| | | | Glutathione-S-transferase, alpha 2/GSTA2 | 138360 | M15872 | 6p12.2 |
| | | | Glutathione-S-transferase, kappa 1/GSTK1 | 602321 | ****** | ****** |
| | | | Glutathione-S-transferase 1/MGST1 (microsomal) | 138330 | AH003674 | Chr.12 |
| | | | Glutathione-S-transferase 2/MGST2 (microsomal) | 601733 | NM_00241 3 | 4q28-q31 |
| | | | Glutathione-S-transferase, mu 1-like/GSTM1L | 138270 | ****** | Chr. 3 |
| | | | 1/GSTM1 | 138350 | J03817 | 1p13.3 |
| | | | Glutathione-S-transferase, mu 2/GSTM2 (muscle) | 138380 | NM_00084 8 | 1p13.3 |
| | | | Glutathione-S-transferase, mu 3/GSTM3 (brain) | 138390 | NM_00084 9 | 1p13.3 |
| | | | Glutathione-S-transferase, mu 4/GSTM4 | 138333 | NM_00085 0 | 1p13.3 |
| | | | Glutathione-S-transferase, mu 5/GSTM5 (brain/lung) | 138385 | NM_00085 1 | 1p13.3 |
| | | | Glutathione-S-transferase, pi/GSTP1 | 134660 | NM_00085 2 | 11q13 |
| | | | Glutathione-S-transferase, theta 1/GSTT1 | 600436 | NM_00085 3 | 22q11.2 |
| | | | Glutathione-S-transferase, theta 2/GSTT2 | 600437 | NM_00085 4 | 22q11.3 |
| | | | Glutathione-S-transferase, zeta 1/maleylacetoactetate isomerase/MAAI/GSTZ1 | 603758 | NM_00151 3 | 14q24.3 |

| | | | | | |
|---|---|---|---|---|---|
| | | | glutathione reductase | 138300 | X15722 | 8p21.1 |
| | | | glutathione peroxidase GPx2 | 138319 | X68314 | 14q24.1: |
| | | | glutathione peroxidase GPx3 | 138321 | X58295 | 5q32-q33.1 |
| | | | glutathione peroxidase GPx1 | 138320 | Y00433 | 3p21.3 |
| | | | glutathione peroxidase GPx4 | 138322 | X71973 | 19p13.3 |
| | | | glutathione peroxidase GPx5 | 603435 | AJ005277 | ****** |
| | | **Metallo-thionein** | metallothionein 2a | 156360 | NM_005953 | 16q13 |
| | | | metallothionein 1g | 156353 | J03910 | 16q13 |
| | | | metallothionein 1f | 156352 | M10943 | 16q13 |
| | | | metallothionein 1e | 156351 | M10942 | 16q13 |
| | | | metallothionein 1b | 156349 | AH001510 | 16q13 |
| | | | metallothionein 3 | 139255 | NM_005954 | 16q13 |
| | | **Proteolysis** | bleomycin hydrolase | 602403 | X92106 | 17q11.2 |
| | | **Methylation** | thiopurine methyltransferase | 187680 | U12387 | 6p22.3 |
| | | **Oxidation** | dehydrogenase/ALDH1 | 100640 | M26761 | 9q21 |
| | | | myeloperoxidase | 254600 | X04876 | 17q23.1 |
| | | **Acetylation** | N-acetyltransferase 1/NAT1 | 108345 | NM_000662 | 8p23.1-p21.3 |
| | | | N-acetyltransferase 2/NAT2 | 243400 | NM_000015 | 8p23.1-p21.3 |
| | | | human DNA mismatch repair protein hMLH1/MutL | 120436 | U07418 | 3p21.3 |
| | | | xeroderma pigmentosum complementation group A/XPA | 278700 | D14533 | 9q22.3 |
| | | | xeroderma pigmentosum complementation group C/XPC | 278720 | NM_004628 | 3p25 |

| | | | |
|---|---|---|---|
| DNA Repair | RAD2 (S. cerevisiae) homolog/RAD2/excision repair complementation group 5/ERCC5 | 133530 | NM_000123 | 13q33 |
| | RAD23 (S. cerevisiae) homolog A/RAD23A | 600061 | NM_005053 | 19p13.2 |
| | RAD23 (S. cerevisiae) homolog B/RAD23B | 600062 | NM_002874 | 3p25.1 |
| | RAD26 (S. cerevisiae) homolog/RAD26/excision repair complementation group 6/ERCC6 | 133540 | NM_000124 | 10q11 |
| | RAD50 (S. cerevisiae) homolog/RAD50 | 604040 | NM_005732 | 5q31 |
| | RAD51 (S. cerevisiae) homolog (E coli RecA homolog)/RAD51 | 179617 | NM_002875 | 15q15.1 |
| | RAD51 (S. cerevisiae) homolog B/RAD51L1 | 602948 | Y15572 | 14q23.3-q24 |
| | RAD51 (S. cerevisiae) homolog D/RAD51L2 | 602774 | NM_002876 | 17q |
| | RAD52 (S. cerevisiae) homolog/RAD52 | 602954 | NM_002878 | 17q11 |
| | RAD51 (S. cerevisiae) homolog C/RAD51L4 | 600392 | NM_002879 | 12p13-q12.2 |
| | RAD54 (S.cerevisiae)-like/RAD54L | 603615 | NM_003579 | 1p32 |
| | excision repair complementation group 1/ERCC1 | 126380 | NM_001983 | 19q13.2-q13.3 |
| | excision repair complementation group 2/ERCC2/XPD | 126340 | L47234 | 19q13.2-q13.3 |
| | excision repair complementation group 1/ERCC3/XPB | 133510 | NM_000122 | 2q21 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **DNA Replication and Repair** | **DNA Replication and Repair** | | excision repair complementation group 4/ERCC4 | 133520 | NM_005236 | 16p13.3-p13.13 |
| | | | replication protein A1 (70kD)/RPA1 | 179835 | NM_002945 | 17p13.3 |
| | | | replication protein A2 (32kD)/RPA2 | 179836 | NM_002946 | 1p35 |
| | | | replication protein A3 (14kD)/RPA3 | 179837 | NM_002947 | 7p22 |
| | | | excision repair protein ERCC1 | 126380 | M13194 | 19q13.2-q13.3 |
| | | | mismatch repair protein hMSH2 | 120435 | U03911 | 2p22-p21 |
| | | | O6 alkylguanine-DNA alkyltransferase | 156569 | M60761 | 10q24.33-qter |
| | | | ADP-ribosyltransferase (NAD+; poly (ADP-ribose) polymerase)/PARP/ADPRT | 173870 | NM_001618 | 1q42 |
| | | | poly (ADP-ribose) glycohydrolase/PARG | 603501 | NM_003631 | 10q11.23 |
| | | | APEX nuclease (multifunctional DNA repair enzyme)/APEX | 107748 | NM_001641 | 14q12 |
| | | | 8-oxoguanine DNA glycosylase/OGG1 | 601982 | NM_002542 | 3p26.2 |
| | | | N-methylpurine-DNA glycosylase/MPG | 156565 | NM_002434 | 16pter-p13.3 |
| | | **Replication** | topoisomerase IIb | 126431 | U54831 | 3p |
| | | | topoisomerase IIa | 126430 | J04088 | 17q21-q22: |
| | | | topoisomerase I | 126420 | J03250 | 20q12-q13.1 |

| | | | | |
|---|---|---|---|---|
| **Mitosis** | beta tubulin 2/TUBB2 | 602660 | NM_006088 | ****** |
| | beta tubulin 4/TUBB4 | 602661 | NM_006086 | ****** |
| | beta tubulin 5/TUBB5 | 602662 | NM_006087 | ****** |
| | gamma tubulin/TUBG | 191135 | NM_001070 | ****** |
| **Histone Acetylation** | histone acetyltransferase | 603053 | AF030424 | 2q31.2-q33.1 |
| | histone deacetylase | 601241 | U50079 | 1p34.1 |
| **Telomere Maintainance** | telomerase protein component 1 | 601686 | U86136 | 14q11.2 |
| | telomerase reverse transcriptase | 187270 | AF015950 | 5p15.33 |
| | telomerase RNA component | 602322 | U86046 | 3q21-q28 |
| **DNA Methylation** | DNA methyltransferase DNMT1 | 126375 | X63692 | 19p13.3-p13.2 |
| | DNA methyltransferase DNMT2 | 602478 | AF012128 | 10p15.1 |
| | DNA methyltransferase DNMT3A | 602769 | AF067972 | 2p23 |
| | DNA methyltransferase DNMT3B | 602900 | NM_006892 | 20q11.2 |
| | thymidylate synthetase | 188350 | X02308 | 18p11.32 |
| | cytidine deaminase | 123920 | L27943 | 1p36.2-p35 |
| | DPD | 274270 | U09178 | 1p22 |
| | deoxycytidine kinase | 125450 | M60527 | 4q13.3-q21.1 |
| | Soluble thymidine kinase 1/TK1 | 188300 | NM_003258 | 17q25.2-q25.3 |
| | 2/TK2 | 188250 | U77088 | Chr.16 |

| | | | | | |
|---|---|---|---|---|---|
| Nucleotide Metabolism | Pyrimidines | uridine kinase | 191730 | NM_003364 | Chr.7 |
| | | uridine monophosphate kinase | 191710 | NM_005267 | 1p32 |
| | | uridine phosphorylase | 191730 | NM_003364 | Chr.7 |
| | | thymidine phosphorylase | 131222 | M58602 | 22q13.32-qter |
| | | aspartate transcarbamylase/CAD trifunctional protein | 114010 | NM_004341 | 2p21 |
| | | orotate phosphoribosyl transferase | 258900 | NM_000373 | 3q13 |
| | Purines | hypoxanthine-guanine phosphoribosyltransferase | 308000 | M31642 | Xq26-q27.2 |
| | | adenosine phosphoribosyltransferase/APRT | 102600 | NM_000485 | 16q24 |
| | | thiopurine S-methyltransferase/TPMT | 187680 | NM_000367 | 6p22.3 |
| | | urate oxidase | 191540 | AH003594 | 1p22 |
| | | adenylosuccinate synthetase/ADSS | 103060 | NM_001126 | 1cen-q12 |
| | | adenylosuccinate lyase | 103050 | NM_000026 | 22q13.1 |
| | | glycinamide ribotide formyltransferase | 138440 | X54199 | 21q22.1 |
| | | purine nucleoside phosphorylase | 164050 | NM_000270 | 14q13.1 |
| | | xanthine oxidase | 278300 | NM_000379 | 2p23-p22 |
| | | adenosine deaminase | 102700 | NM_000022 | 20q13.11 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Cellular Metabolism | | Ribo-nucleotides | ribonucleotide reductase M1 subunit | 180410 | L10342 | 11p15.5 |
| | | | ribonucleotide reductase M2 subunit | 180390 | X59618 | 2p25-p24 |
| | | General | ecto-5'-nucleotidase (CD73)/NT5 | 129190 | X55740 | 6q14-q21 |
| | | | alkaline phosphatase | 171760 | NM_000478 | 1p36.1-p34 |
| | Amino Acid Metabolism | Amino Acid Metabolism | asparagine synthetase | 108370 | M27396 | 7q21-q31 |
| | | | arginase (ARG1) | 207800 | M14502 | ****** |
| | | | arginase (ARG2) | 107830 | D86724 | 14q24.1-q24.3 |
| | | | ornithine transaminase | 258870 | NM_000274 | 10q26 |
| | Steroid Metabolism *(additional genes in Endocrinology and Metabolism)* | Steroid Metabolism | cytochrome P450 aromatase (CYP19) | 107910 | X13589 | 15q21.1 |
| | | | steroid 5 alpha reductase | 184753 | M32313 | 5p15 |
| | | | estrogen sulfotransferase | 600043 | U20521 | 4q13.1 |
| | | | steroid 5-alpha reductase 2 | 264600 | NM_000348 | 2p23 |
| | | | HMGCoA reductase | 142910 | NM_000859 | 5q13.3-q14 |
| | | | squalene synthetase | 184420 | X69141 | 8p23.1-p22 |
| | | | Folate Receptor Alpha/FOLR1 | 136430 | M28099 | 1q13.3-q13.3 |
| | | | Folate Receptor Beta/FOLR2 | 136425 | AF000380 | 1q13.3-q13.3 |
| | | | Folate Receptor Gamma/FOLR3 | 602469 | Z32564 | ****** |
| | | | Folate Transporter (SLC19A1) | 600424 | U19720 | 21q22.3 |
| | | | Vitamin B12 binding protein | 275350 | NM_000355 | 22q11.2-qter |
| | | | folylpolyglutamate synthetase/FPGS | 136510 | M98045 | 9cen-q34 |
| | | | gamma-glutamyl hydrolase/GGH | 601509 | U55206 | ****** |
| | | | Methylenetetrahydrofolate reductase/MTHFR | 236250 | U09806 | 1p36.3 |
| | | | Dihydrofolate reductase/DHFR | 126060 | J00140 | 5q11.2-q13.2 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | **Folate Metabolism** | **Folate Metabolism** | 5,10-methylenetetrahydrofolate dehydrogenase, 5,10-methylenetetrahydrofolate cyclohydrolase, 10-formyltetrahydrofolate synthetase/MTHFD1 | 172460 | NM_005956 | 14q24 |
| | | | 5,10-methenyltetrahydrofolate synthetase (5-formyltetrahydrofolate cyclo-ligase)/MTHFS | 604197 | NM_006441 | Chr. 15 |
| | | | phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole | 138440 | NM_000819 | 21q22.1 |
| | | | folate hydrolase 1/FOH1 | 600934 | NP_004467 | 11q14 |
| | | | 6-pyruvoyl tetrahydrobiopterin synthase/PTPS | 261640 | Q03393 | 1q22.3-q23.3 |
| | | | serine hydroxymethyltransferase 1 (soluble)/SHMT1 | 182144 | NM_004169 | 17p11.2 |
| | | | serine hydroxymethyltransferase 2 (mitochondrial)/SHMT2 | 138450 | NM_005412 | 12q13 |
| | | | Glycine aminotransferase/glycine cleavage T protein/GAT | 238310 | NM_000481 | 3p21.2-p21.1 |
| | | | 5-methyltetrahydrofolate-homocysteine methyltransferase/methionine | 156570 | NM_000254 | 1q43 |
| | | | glutamate formiminotransferase/dihydrofolate synthetase | 229100 | ****** | ****** |

| | | | | | | |
|---|---|---|---|---|---|---|
| | **Protein Modification** | **Prenylation** | farnesyl:protein transferase alpha/FNTA | 134635 | NM_002027 | 8p22-q11 |
| | | | farnesyl:protein transferase beta/FNTB | 134636 | NM_005023 | 14q23-q24 |
| | | | Rab geranylgeranyltransferase, alpha subunit/RABGGTA | 601905 | NM_004581 | 14q11.2 |
| | | | Rab geranylgeranyltransferase, beta subunit/RABGGTB | 179080 | NM_004582 | 1p31-p22 |
| | **Polyamine Metabolism** | **Polyamine Metabolism** | ornithine decarboxylase 1 (ODC1) | 165640 | M16650 | 2p25 |
| | | | SAM decarboxylase | 180980 | M21154 | 6q21-q22 |
| | **Phospholipid Metabolism** | **Phospholipid Metabolism** | glucosylceramide synthase | 602874 | D50840 | 9q31 |
| | | | interferon alpha1 (IFNa1) | 147660 | X02956 | 9p22 |
| | | | interferon alpha2 (IFNa2) | 147562 | ****** | 9p22 |
| | | | interferon beta1 (IFNb1) | 147640 | V00546 | 9p21 |
| | | | interferon beta3 (IFNb3) | 147860 | ****** | Chr.8 |
| | | | interferon omega1 (IFNw1) | 147553 | X02669 | 9p21 |
| | | | interferon gamma (IFNg) | 147570 | L07633 | 12q14 |
| | | | interleukin 1 alpha (IL1a) | 147760 | M15329 | 2q14 |
| | | | interleukin 1 beta (IL1b) | 147720 | K02770 | 2q14 |
| | | | interleukin 2 (IL2) | 147680 | X01586 | 4q26-q27 |
| | | | interleukin 3 (IL3) | 147740 | M20137 | 5q31.1 |
| | | | interleukin 4 (IL4) | 147780 | M13982 | 5q31.1 |
| | | | interleukin 5 (IL5) | 147850 | X04688 | 5q31.1 |
| | | | interleukin 6 (IL6) | 147620 | M14584 | 7p21 |
| | | | interleukin 7 (IL7) | 146660 | J04156 | 8q12-q13 |
| | | | interleukin 8 (IL8) | 146930 | M26383 | 4q12-q13 |
| | | | interleukin 9 (IL9) | 146931 | X17543 | 5q31.1 |
| | | | interleukin 10 (IL10) | 124092 | M57627 | 1q31-q32 |

| | | | |
|---|---|---|---|
| Cytokines | interleukin 11 (IL11) | 147681 | X58377 | 19q13.3-q13.4 |
| | interleukin 12a (IL12a) | 161560 | NM_002187 | 3p12-q13.2 |
| | interleukin 12b (IL12b) | 161561 | NM_000440 | 5q31.1-q33.1 |
| | interleukin 13 (IL13) | 147683 | X69079 | 5q31 |
| | interleukin 15 (IL15) | 600554 | U14407 | 4q31 |
| | interleukin 16 (IL16) | 603035 | NM_004513 | ****** |
| | interleukin 18 (IL18) | 600953 | ****** | 11q22.2-q22.3 |
| | interferon alpha receptor 1 (IFNAR1) | 107450 | X77722 | 21q22.1 |
| | interferon alpha receptor 2 (IFNAR2) | 147569 | U68755 | 21q22.1-q22.2 |
| | (IFNGR1) | 107470 | J03143 | 6q23-q24 |
| | interferon gamma receptor 2 (IFNGR2) | 602376 | NM_000874 | 21q22.1 |
| | interleukin 1 receptor 1 (IL-1R1) | 147810 | M27492 | 2q12 |
| | interleukin 1 receptor 2 (IL-1R2) | 147811 | NM_004633 | 2q12-q22 |
| | interleukin 2 receptor alpha (IL-2Ra) | 147730 | X01057 | 10p15-p14 |
| | interleukin 2 receptor beta (IL-2Rb) | 146710 | M26062 | 22q11.2-q13 |
| | interleukin 2 receptor gamma (IL-2Rg) | 308380 | D11086 | Xq13 |
| | interleukin 3 alpha receptor (IL-3aR) | 308385 | M74782 | Xp22.3 |
| | interleukin 4 receptor (IL-4R) | 147781 | X52425 | 16p12.1-p11.2 |
| | interleukin 5 receptor alpha (IL-5Ra) | 147851 | M96652 | 3p26-p24: |
| | interleukin 6 receptor (IL-6R) (20) | 147880 | X12830 | 1q21 |

| | | | | |
|---|---|---|---|---|
| **Inflammation** *(additional genes in Immunology)* | interleukin 7 receptor (IL-7R) | 146661 | M29696 | 5p13 |
| | interleukin 8 receptor alpha (IL-8Ra) | 146929 | M68932 | 2q35 |
| | interleukin 8 receptor beta (IL-8Rb) | 146928 | M94582 | 2q35 |
| | interleukin 9 receptor (IL-9R) | 300007 | M84747 | Xq28 |
| | 10Ra) | 146933 | U00672 | 11q23.3 |
| | interleukin receptor 11 alpha (IL-11a) | 600939 | U32324 | 9p13 |
| | interleukin receptor 12 beta (IL-12b) | 600939 | U03187 | 9p13 |
| | 12b2) | 601642 | U03187 | 1p31.2 |
| | interleukin receptor 13 alpha (IL-13a) | 300119 | S80963 | Chr.X |
| | 13a2) | 300130 | X95302 | Xq24 |
| | 15Ra) | 601070 | U31628 | 10p15-p14 |
| | tumor necrosis factor alpha/TNFA | 191160 | X01394 | 6p21.3 |
| | tumor necrosis factor beta/TNFB/lymphotoxin alpha/LTA | 153440 | NM_000595 | 6p21.3 |
| | tumor necrosis factor ligand superfamily, member 4/TNFSF4 | 603594 | NM_003326 | 1q25 |
| | tumor necrosis factor ligand superfamily, member 5/TNFSF5 | 308230 | NM_000074 | Xq26 |
| | tumor necrosis factor ligand superfamily, member 6/TNFSF6 | 134638 | ****** | 1q23 |
| | tumor necrosis factor ligand superfamily, member 7/TNFSF7 | 602840 | NM_001252 | 19p13 |
| | tumor necrosis factor ligand superfamily, member 8/TNFSF8 | 603875 | NM_001244 | 9q33 |
| | tumor necrosis factor ligand superfamily, member 10/TNFSF10 | 603598 | NM_003810 | 3q26 |
| | tumor necrosis factor ligand superfamily, member 11/TNFSF11 | 602642 | NM_003701 | 13q14 |
| | tumor necrosis factor ligand superfamily, member 12/TNFSF12 | 602695 | NM_003809 | 17p13.3 |

| | | | | |
|---|---|---|---|---|
| **Inflammation (additional genes in Immunology)** | tumor necrosis factor ligand superfamily, member 13B/TNFSF13B | 603969 | NM_006573 | 13q32-q34 |
| | tumor necrosis factor ligand superfamily, member 15/TNFSF15 | 604052 | ****** | 9q33 |
| | macrophage inflammatory protein 1 alpha | 182283 | M23178 | 17q11-q21 |
| | tumor necrosis factor ligand superfamily, member 18/TNFSF18 | 603898 | ****** | 1q23 |
| | myeloid progenitor inhibitory factor 1 | 602494 | ****** | ****** |
| | macrophage inflammatory protein 1 alp | 182283 | M23178 | 17q11-q21 |
| **Interferon Response** | 2',5'-oligoadenylate synthetase 1 (OAS1) | 164350 | NM_006187 | 12q24.2 |
| | 2',5'-oligoadenylate synthetase 2 (OAS2) | 603350 | M87284 | 12q24.2 |
| | 2',5'-oligoadenylate synthetase 3 (OAS3) | 603351 | ****** | 12q24.2 |
| | arachidonate 5' lipoxygenase/ALOX5 | 152390 | J03571 | Chr.10 |
| | arachidonate 12-lipoxygenase/ALOX12 | 152391 | NM_000697 | 17p13.1 |
| | prostaglandin endoperoxide synthetase 1/COX1/PTGS1 | 176805 | AH001520 | 9q32-q33.3 |
| | prostaglandin endoperoxide synthetase 2/COX2/PTGS2 | 600262 | NM_000963 | 1q25.2-q25.3 |
| | thromboxane A synthase 1/TBXAS1 | 274180 | SEG_D34613S | 7q34 |
| | prostaglandin D2 synthase | 602598 | M61900 | ****** |
| | prostaglandin I2 synthase/prostacyclin synthase/PTGIS | 601699 | SEG_D83393S | 20q13 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | Prostaglandin | prostaglandin E receptor 1, EP1 subtype/PTGER1 | 176802 | NM_000955 | 19p13.1 |
| | | | prostaglandin E receptor 2, EP2 subtype/PTGER2 | 176804 | ****** | 5p13.1 |
| | | | prostaglandin E receptor 3, EP3 subtype/PTGER3 | 176806 | NM_000957 | 1p31.2 |
| | | | prostaglandin E receptor 4, EP4 subtype/PTGER4 | 601586 | NM_000958 | 5p13.1 |
| | | | prostaglandin F receptor/PTGFR | 600563 | L24470 | 1p31.1 |
| | | | prostaglandin F2 receptor negative regulator/PTGFRN | 601204 | U26664 | 1p13.1-q21.3 |
| | | | prostaglandin I2 receptor/PTGIR/prostacyclin receptor | 600022 | SEG_HUMIP | 19q13.3 |
| | | | 15-hydroxyprostaglandin dehydrogenase/HPGD | 601688 | NM_000860 | 4q34-q35 |
| | | | aldo-keto reductase family 1, member C2/AKR1C2 | 600450 | NM_001353 | 10p15-p14 |
| | | | integrin alpha 1 | 192968 | Y00796 | Chr.5 |
| | | | integrin alpha 2 | 192974 | X17033 | 5q23-q31 |
| | | | integrin alpha 4 | 192975 | L12002 | 2q31-q32 |
| | | | integrin alpha 5 | 135620 | NM_002205 | 12q11-q13 |
| | | | integrin alpha 6 | 147556 | X59512 | Chr.2 |
| | | | integrin alpha 7 | 600536 | AF032108 | 12q13 |
| | | | integrin alpha 8 | 604063 | L36531 | ****** |
| | | | integrin alpha 9 | 603963 | L24158 | ****** |
| | | | integrin alpha 10 | 604042 | AF074015 | ****** |
| | | | integrin alpha D | 602453 | U40279 | 16p11.2 |
| | | | integrin alpha M | 120980 | J04145 | 16p11.2 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Adhesion | Integrins | integrin alpha X | 151510 | M81695 | 16p11.2 |
| | | | integrin beta 1 | 135630 | U28252 | 10p11.2 |
| | | | integrin beta 2 | 600065 | M15395 | 21q22.3 |
| | | | integrin beta 3 | 173470 | M35999 | 17q21.32 |
| | | | integrin beta 4 | 147557 | X51841 | 17q11-qter |
| | | | integrin beta 5 | 147561 | M35011 | ****** |
| | | | integrin beta 6 | 147558 | M35198 | Chr.2 |
| | | | integrin beta 7 | 147559 | M68892 | 12q13.13 |
| | | | integrin beta 8 | 604160 | NM_002214 | ****** |
| | | | cadherin 2/NCAD/CDH2 | 114020 | Z27440 | 18q11.2 |
| | | | human cell adhesion protein SQM1 | 603842 | NM_004146 | 19p13.12-p13.11 |
| | | Proteases | matrix metalloproteinase 3, stromelysin 1 | 185250 | NM_002422 | 11q23 |
| | | | matrix metalloproteinase 1, | 120353 | M13509 | 11q22-q23 |
| | | | aminopeptidase A/glutamyl aminopeptidase/ENPEP | 138297 | L14721 | 4q25 |
| | | | mammary serine protease/protease M/neurosin | 602652 | D78203 | 19q13.3 |
| | | | protease inhibitor 5/maspin | 154790 | NM_002639 | 18q21.3 |
| | | | activator | 191840 | AH007073 | 10q24 |
| | | | cathepsin B | 116810 | M14221 | 8p22 |
| | | | type 2 plasminogen activator inhibitor | 173390 | NM_002575 | 18q21.3 |
| | | | urokinase-type plasminogen activator receptor | 173391 | NM_002659 | 19q13 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | **DNA Damage Checkpoint** | RAD1 (S. pombe) homolog/RAD1 | 603153 | NM_002853 | 5p13.3-p13.2 |
| | | | RAD9 (S. pombe) homolog/RAD9 | 603761 | NM_004584 | 11q13.1-q13.2 |
| | | | RAD17 (S. pombe) homolog/RAD17 | 603139 | NM_002873 | 4q13.3-q21.2 |
| | | | FRAP-related protein/FRP1/ATR | 601215 | U49844 | 3q22-q24 |
| | | | HUS1 (S. pombe) checkpoint homolog/HUS1 | 603760 | NM_004507 | 7p13-p12 |
| | | | ataxia telangiectasia mutated (complementation groups A, C and D)/ATM | 208900 | NM_000051 | 11q22.3 |
| | | | CHK1 (checkpoint, S.pombe) homolog/CHEK1 | 603078 | NM_001274 | 11q22-q23 |
| | | | growth arrest and DNA-damage-inducible, alpha/GADD45A | 126335 | NM_001924 | 1p34-p12 |
| | | | BRCA1 | 113705 | NM_000058 | 17q21 |
| | | | BRCA2 | 600185 | NM_000059 | 13q12.3 |
| | | | benzimidazoles 1 (yeast homolog)/BUB1 | 602452 | AF139363 | 2q12-q14 |
| | | | benzimidazoles 1 (yeast homolog), beta/BUB1B | 602860 | NM_001211 | 15q14-q21 |
| | | | benzimidazoles 1 (yeast homolog)/BUB3 | 603719 | NM_004725 | 10q24-q26 |
| | | | MAD (mothers against decapentaplegic) homolog 1/MADH1 | 601595 | NM_005900 | 4q28 |
| | | | MAD (mothers against decapentaplegic) homolog 2/MADH2 | 601366 | NM_005901 | 18q21 |

| | | | |
|---|---|---|---|
| MAD (mothers against decapentaplegic) homolog 3/MADH3 | 603109 | NM_005902 | 15q21-q22 |
| MAD (mothers against decapentaplegic) homolog 4/MADH4 | 600993 | NM_005359 | 18q21.1 |
| MAD (mothers against decapentaplegic) homolog 5/MADH5 | 603110 | NM_005903 | 5q31 |
| MAD (mothers against decapentaplegic) homolog 6/MADH6 | 602931 | NM_005585 | 15q21-q22 |
| MAD (mothers against decapentaplegic) homolog 7/MADH7 | 602932 | NM_005904 | Chr.18 |
| MAD (mothers against decapentaplegic) homolog 9/MADH9 | 603295 | NM_005905 | 13q12-q14 |
| cyclin-dependent kinase (CDK2) | 116953 | NM_001798 | 12q13 |
| cyclin-dependent kinase (CDK3) | 123828 | NM_001258 | 17q22-qter |
| cyclin-dependent kinase (CDK4) | 123829 | NM_000075 | 12q14 |
| cyclin-dependent kinase (CDK5) | 123831 | NM_004935 | 7q36 |
| cyclin-dependent kinase (CDK6) | 603368 | NM_001259 | 7q21-q22 |
| cyclin-dependent kinase (CDK7) | 601955 | NM_003157 | 2p15-cen |
| cyclin-dependent kinase (CDK8) | 603184 | NM_001260 | 13q12 |
| cyclin-dependent kinase (CDK9) | 603251 | NM_001261 | 9q34.1 |
| cyclin A1 | 604036 | U66838 | ****** |
| cyclin A2 | 123835 | X51688 | 4q27 |

| | | | | |
|---|---|---|---|---|
| Cell Cycle | cyclin B1 | 123836 | M25753 | 5q12 |
| | cyclin B2 | 602755 | AF002822 | ****** |
| | cyclin C | 123838 | M74091 | 6q21 |
| | cyclin D1 | 168461 | M73554 | 11q13 |
| | cyclin D2 | 123833 | M90813 | 12p13 |
| | cyclin D3 | 123834 | M90814 | 6p21 |
| | cyclin E1 | 123837 | U40739 | 19q13.1 |
| | cyclin E2 | 603775 | AF091433 | ****** |
| | cyclin F | 600227 | Z36714 | 16p13.3 |
| | cyclin G1 | 601578 | X77794 | 5q32-q34 |
| | cyclin G2 | 603203 | U47414 | ****** |
| | cyclin H | 601953 | U11791 | 5q13.3-q14 |
| | cyclin K | 603544 | AF060515 | 14q32 |
| | cyclin T1 | 602506 | AF045161 | Chr.12 |
| | cyclin T2 | 603862 | AF048731 | ****** |
| | cyclin-dependent kinase inhibitor 1a/WAF | 116899 | U03106 | 6p21.2 |
| | cyclin-dependent kinase inhibitor 1b/KIP1 | 600778 | U10906 | 12p13 |
| | cyclin-dependent kinase inhibitor 2a | 600160 | NM_000077 | 9p21 |
| | cell cycle CDC2 | 116940 | NM_001786 | 10q21.1 |
| | cell division cycle 25A/CDC25A | 116947 | NM_001789 | 3p21 |
| | E2F transcription factor 1/E2F1 | 189971 | M96577 | 20q11.2 |
| | E2F transcription factor 2/E2F2 | 600426 | NM_004091 | 1p36 |
| | E2F transcription factor 3/E2F3 | 600427 | NM_001949 | 6p22 |

| | | | |
|---|---|---|---|
| E2F transcription factor 4/E2F4 | 600659 | NM_001950 | 16q22.1 |
| E2F transcription factor 5/E2F5 | 600967 | NM_001951 | ****** |
| E2F transcription factor 6/E2F6 | 602944 | NM_001952 | ****** |
| transcription factor Dp-1 (E2F dimerization partner 1)/TFDP1 | 189902 | NM_007111 | 13q34 |
| transcription factor Dp-2 (E2F dimerization partner 2)/TFDP2 | 602160 | NM_006286 | 3q23 |
| retinoblastoma-related gene RB2/p130 | 180203 | NM_005611 | 16q12.2 |
| retinoblastoma RB1 | 180200 | NM_000321 | 13q14.1-q14.2 |
| insulin-like growth factor 1/somatomedin C/IGF1 | 147440 | M11568 | 12q22-q24.1 |
| insulin-like growth factor 2/somatomedin A/IGF2 | 147470 | NM_000612 | 11p15.5 |
| insulin-like growth factor binding protein 1 | 146730 | NM_000596 | 7p14-p12 |
| insulin-like growth factor binding protein 2 | 146731 | X16302 | 2q33-q34 |
| insulin-like growth factor binding protein 3 | 146732 | NM_000598 | 7p14-p12 |
| insulin-like growth factor binding protein 4 | 146733 | M62403 | 17q12-q21 |
| insulin-like growth factor binding protein 5 | 146734 | L27560 | ****** |
| insulin-like growth factor binding protein 6 | 146735 | M69054 | Chr.12 |

| | | | | |
|---|---|---|---|---|
| Growth Factors | insulin-like growth factor binding protein 7 | 602867 | L19182 | 4q12 |
| | insulin-like growth factor binding protein 10 | 602369 | U62015 | 1p22.3 |
| | schwannoma-derived growth factor/amphiregulin/AREG | 104640 | NM_001657 | 4q13-q21 |
| | trefoil factor 1/sP2 | 113710 | X00474 | 21q22.3 |
| | trefoil factor 2/TFF2 | 182590 | X51698 | 21q22.3 |
| | trefoil factor 3/sP2 | 600633 | L08044 | 21q22.3 |
| | epidermal growth factor EGF | 131530 | X04571 | 4q25 |
| | transforming growth factor (TGF-B1) | 190180 | M60315 | 19q13.1-q13.3 |
| | transforming growth factor (TGF-B2) | 190220 | M19154 | 1q41 |
| | transforming growth factor (TGF-B3) | 190230 | X14149 | 14q24 |
| | vascular endothelial growth factor (VEGF-A) | 192240 | M32977 | 6p12 |
| | vascular endothelial growth factor (VEGF-B) | 601398 | U52819 | 11q13 |
| | vascular endothelial growth factor (VEGF-C) | 601528 | X94216 | ****** |
| | erythropoietin/EPO | 133170 | NM_000799 | 7q21 |
| | cardiotrophin 1 | 600435 | ****** | ****** |
| | leukemia inhibitory factor/LIF | 159540 | NM_002309 | 22q12.1-q12.2 |
| | ciliary neurotrophic factor/CNTF | 118945 | NM_000614 | 11q12.2 |
| | oncostatin M | 165095 | ****** | 22q12.1-q12.2 |
| | heparin-binding growth factor 1/FGF1 | 131220 | AH002717 | 5q31 |

| | | | |
|---|---|---|---|
| basic fibroblast growth factor/FGF2 | 134920 | M27968 | 4q25-q27 |
| fibroblast growth factor 3/FGF3 | 164950 | NM_005247 | 11q13 |
| HST oncogene/fibroblast growth factor 4/FGF4 | 164980 | NM_002008 | 11q13 |
| fibroblast growth factor-related protein/FGF5 | 165190 | AH005423 | 4q21 |
| fibroblast growth factor 6/FGF6 | 134921 | X63454 | 12p13 |
| keratinocyte growth factor/fibroblast growth factor 7/FGF7 | 148180 | L06245 | 15q15-q21.1 |
| fibroblast growth factor 8 (androgen-induced)/FGF8 | 600483 | NM_006119 | 10q24 |
| fibroblast growth factor 9 (glia-activating factor)/FGF9 | 600921 | NM_002010 | 13q11-q12 |
| fibroblast growth factor 10/FGF10 | 602115 | NM_004465 | 5p13-p12 |
| fibroblast growth factor 11/FGF11 | 601514 | NM_004112 | 17q21 |
| fibroblast growth factor 12/FGF12 | 601513 | ****** | 3q28 |
| fibroblast growth factor 13/FGF13 | 300070 | ****** | Xq26.3 |
| fibroblast growth factor 14/FGF14 | 601515 | NM_004115 | 13q34 |
| fibroblast growth factor 16/FGF16 | 603724 | NM_003868 | ****** |
| fibroblast growth factor 17/FGF17 | 603725 | NM_003867 | 8p21 |
| fibroblast growth factor 18/FGF18 | 603726 | NM_003862 | ****** |
| fibroblast growth factor 19/FGF19 | 603891 | NM_005117 | ****** |

| | | | |
|---|---|---|---|
| stem cell factor/mast cell growth factor (MGF) | 184745 | NM_003994 | 12q22 |
| thrombopoietin/THPO | 600044 | NM_000460 | 3q26.3-q27 |
| platelet derived growth factor beta polypeptide/PDGFB/SIS | 190040 | NM_002608 | 22q12.3-q13.1 |
| insulin-like growth factor 1 receptor precursor/IGF1R | 147370 | NM_000875 | 15q25-q26 |
| insulin-like growth factor 2 receptor/IGF2R | 147280 | NM_000876 | 6q26 |
| epidermal growth factor receptor EGFR | 131550 | NM_005228 | 7p12.3-p12.1 |
| tyrosine kinase-type cell surface receptor HER2/ERBB2/NEU | 164870 | X03363 | 17q21.2 |
| glucocorticoid receptor | 138040 | M11050 | 5q31 |
| glucocorticoid receptor alpha | 138040 | U25029 | 5q31 |
| glucocorticoid receptor beta | 138040 | X03348 | 5q31 |
| progesterone receptor | 264080 | M15716 | 11q22 |
| androgen receptor | 313700 | M20132 | Xq11-q12 |
| estrogen receptor 1 (ESR1) | 133430 | M12674 | 6q25.1 |
| estrogen receptor 2 (ESR2) | 601663 | X99101 | 14q |
| retinoic acid receptor alpha (RARA) | 180240 | X06538 | 17q12 |
| retinoic acid receptor beta (RARB) | 180220 | X07282 | 3p24 |
| retinoic acid receptor gamma (RARG) | 180190 | M38258 | 12q13 |
| Peroxisome proliferative activated receptor, alpha/PPARA | 170998 | NM_005036 | 22q12-q13.1 |
| Peroxisome proliferative activated receptor, gamma/PPARG | 601487 | NM_005037 | 3p25 |
| Peroxisome proliferative activated receptor, delta/PPARD | 180231 | NM_006238 | 1q21.3 |

| | | | |
|---|---|---|---|
| c-kit (MGF receptor) | 164920 | X06182 | 4q12 |
| TGF-B type I receptor | 190181 | AH006005 | 9q33-q34 |
| TGF-B type II receptor | 190182 | NM_003242 | 3p22 |
| TGF-B type III receptor | 600742 | L07594 | 1p33-p32 |
| fibroblast growth factor receptor 1/FGFR1 | 136350 | ****** | 8p11.2-p11.1 |
| fibroblast growth factor receptor 2/FGFR2 | 176943 | Y17131 | 10q26 |
| fibroblast growth factor receptor 3/FGFR3 | 134934 | NM_005247 | 4p16.3 |
| fibroblast growth factor receptor 4/FGFR4 | 134935 | NM_002011 | 5q35.1-qter |
| VEGF receptor | 191306 | X61656 | 4q12 |
| mitogen activated protein kinase PRKM1/MAPK1 | 176948 | NM_002745 | 22q11.2 |
| mitogen activated protein kinase PRKM3/MAPK3 | 601795 | X60188 | 16p11.2 |
| mitogen activated protein kinase PRKM4/MAPK4 | 176949 | NM_002747 | 18q12-q21 |
| mitogen activated protein kinase PRKM6/MAPK6 | 602904 | NM_002748 | ****** |
| mitogen activated protein kinase PRKM7/MAPK7 | 602521 | NM_002749 | 17p11.2 |
| mitogen activated protein kinase JNK1/PRKM8/MAPK8 | 601158 | L26318 | ****** |
| mitogen activated protein kinase JNK2/PRKM9/MAPK9 | 602896 | U09759 | 5q35 |
| mitogen activated protein kinase JNK3/PRKM10/MAPK10 | 602897 | U35003 | ****** |

| | | | | |
|---|---|---|---|---|
| | | mitogen activated protein kinase PRKM11/MAPK11 | 602898 | AF031135 | ****** |
| | | mitogen activated protein kinase SAPK3/MAPK12 | 602399 | NM_002969 | 22q13.3 |
| | | mitogen activated protein kinase PRKM13/MAPK13 | 602899 | NM_002754 | ****** |
| | | mitogen activated protein kinase SAPK2A/MAPK14 | 600289 | NM_001315 | 6p21.3-p21.2 |
| | | vitamin D3 receptor | 601769 | NM_000376 | 12q12-q14 |
| | | transferrin receptor | 190010 | NM_003234 | 3q29 |
| | | thyroid stimulating hormone receptor | 603372 | NM_000369 | 14q31 |
| | **Receptors** | TEK tyrosine kinase receptor (TIE-2) | 600221 | X60957 | 9p21 |
| | | totipotent stem cell receptor FLK 2 | 600007 | U03858 | ****** |
| | | leutenizing hormone choriogonadotropin receptor/LHCGR | 152790 | ****** | 2p21 |
| | | vitamin B12 receptor/cubilin/CUBN | 602997 | NM_001081 | 10p12.1 |
| | | neurotrophic tyrosine kinase receptor/NTRK1/TRKA | 191315 | X03541 | 1q21-q22 |
| | | neurotrophic tyrosine kinase receptor/NTRK2/TRKB | 600456 | X75958 | 9q22.1 |
| | | neurotrophic tyrosine kinase receptor/NTRK3/TRKC | 191316 | U05012 | 15q25 |
| | | colony stimulating factor 1 receptor/CSFR1 | 164770 | U63963 | 5q33.2-q33.3 |

| | | | | | |
|---|---|---|---|---|---|
| **Cellular Growth and Differentiation** | **Cellular Growth and Differentiation** | granulocyte-macrophage colony stimulating factor 2 receptor, alpha, low-affinity/CSF2RA | 306250 | NM_006140 | Xp22.32 |
| | | granulocyte-macrophage colony stimulating factor 2 receptor, beta/CSF2RB | 138981 | U18373 | 22q12.2-q13.1 |
| | | granulocyte-macrophage colony stimulating factor 2 receptor, alpha, Y chromasomal/CSF2RY | 425000 | ****** | Yp11 |
| | | MTV oncogene homolog 1/AKT1 | 164730 | K02777 | 14q32.3 |
| | | MTV oncogene homolog 2/AKT2 | 164731 | NM_001626 | 19q13.1-q13.2 |
| | | erythropoietin receptor/EPOR | 133171 | NM_000121 | 19p13.3-p13.2 |
| | | neutrophil chemotactic response receptor/gp130 | 162820 | ****** | 7q22-qter |
| | | ciliary neurotrophic factor receptor/CNTFR | 118946 | NM_001842 | 9p13 |
| | | tumor necrosis factor receptor superfamily, member 1A/TNFRSF1A | 191190 | NM_001065 | 12p13.2 |
| | | tumor necrosis factor receptor superfamily, member 1B/TNFRSF1B | 191191 | NM_001066 | 1p36.3-p36.2 |
| | | tumor necrosis factor receptor superfamily, member 4/TNFRSF4 | 600315 | NM_003327 | 1p36 |
| | | tumor necrosis factor receptor superfamily, member 5/TNFRSF5 | 109535 | NM_001250 | 20q12-q13.2 |
| | | tumor necrosis factor receptor superfamily, member 6B/TNFRSF6B | 603361 | NM_003823 | 20q13 |
| | | tumor necrosis factor receptor superfamily, member 7/TNFRSF7 | 186711 | ****** | 12p13 |

| | | | |
|---|---|---|---|
| tumor necrosis factor receptor superfamily, member 8/TNFRSF8 | 153243 | NM_001243 | 1p36 |
| tumor necrosis factor receptor superfamily, member 9/TNFRSF9 | 602250 | NM_001561 | 1p36 |
| superfamily, member 10A/TNFRSF10A | 603611 | NM_003844 | 8p21 |
| superfamily, member 10B/TNFRSF10B | 603612 | NM_003842 | 8p22-p21 |
| superfamily, member 10C/TNFRSF10C | 603613 | AF014794 | 8p22-p21 |
| superfamily, member 10D/TNFRSF10D | 603614 | NM_003840 | 8p21 |
| superfamily, member 11A/TNFRSF11A | 603499 | NM_003839 | 18q22.1 |
| superfamily, member 11B/TNFRSF11B | 602643 | NM_002546 | 8q24 |
| tumor necrosis factor receptor superfamily, member 12/TNFRSF12 | 603366 | NM_003790 | 1p36.3 |
| tumor necrosis factor receptor superfamily, member 14/TNFRSF14 | 602746 | NM_003820 | 1p36.3-p36.2 |
| tumor necrosis factor receptor superfamily, member 16/TNFRSF16 | 162010 | NM_002507 | 17q21-q22 |
| tumor necrosis factor receptor superfamily, member 17/TNFRSF17 | 109545 | Z14954 | 16p13.1 |
| tumor necrosis factor receptor superfamily, member 18/TNFRSF18 | 603905 | ****** | 1p36.3 |
| proliferating cell nuclear antigen | 176740 | J04718 | 20p12 |
| protein kinase C alpha | 176960 | X52479 | 17q22-q23.2 |
| protein kinase C beta | 176970 | X06318 | 16p11.2 |

| | | | |
|---|---|---|---|
| protein kinase C delta | 176977 | L07861 | 3p |
| protein kinase C gamma | 176980 | ****** | 19q13.4 |
| protein kinase C theta | 600448 | L01087 | 10p15 |
| protein kinase C zeta | 176982 | L14283 | ****** |
| casein kinase 1 alpha 1 | 600505 | NM_001892 | 13q13 |
| casein kinase 1 gamma 2 | 602214 | U89896 | 19p13.3 |
| casein kinase 1 delta | 600864 | NM_001893 | 17q25 |
| casein kinase 1 epsilon | 600863 | NM_001894 | 22q12-q13 |
| casein kinase 2 alpha 1 | 115440 | J02853 | 20p13 |
| casein kinase 2 alpha 2 | 115442 | NM_001896 | 16p13.3-p13.2 |
| casein kinase 2 beta | 115441 | X57152 | 6p21.3 |
| mitogen-activated protein kinase kinase 1/MAP2K1 | 176872 | NM_002755 | 15q22.1-q22.33 |
| mitogen-activated protein kinase kinase 2/MAP2K2 | 601263 | L11285 | ****** |
| mitogen-activated protein kinase kinase 3/MAP2K3 | 602315 | NM_002756 | 17q11.2 |
| mitogen-activated protein kinase kinase 4/MAP2K4 | 601335 | NM_003010 | 17p11.2 |
| mitogen-activated protein kinase kinase 5/MAP2K5 | 602520 | NM_002757 | ****** |
| mitogen-activated protein kinase kinase 6/MAP2K6 | 601254 | U39065 | ****** |
| mitogen-activated protein kinase kinase 7/MAP2K7 | 603014 | NM_005043 | ****** |
| fos proto-oncogene/FOS | 164810 | K00650 | 14q24.3 |

| | | | |
|---|---|---|---|
| myc proto-oncogene/MYC | 190080 | V00568 | 8q24.12-q24.13 |
| clustrin/TRPM-2 | 185430 | M64722 | 8p21-p12 |
| c-jun | 165160 | J04111 | 1p32-p31 |
| c-myb | 189990 | M15024 | 6q22 |
| mdm-2 | 164785 | Z12020 | 12q14.3-q15 |
| NF kappaB | 164012 | ****** | 10q24 |
| cAMP-dependent protein kinase/protein kinase A | 176910 | X14968 | 3p21.3-p21.2 |
| raf-1 | 164760 | NM_002880 | 3p25 |
| H-ras | 190020 | J00277 | 11p15.5 |
| K-ras | 190070 | K01912 | 12p12.1 |
| N-ras | 164790 | X02751 | 1p13.2 |
| 1/WT1 | 194070 | AH003034 | 11p13 |
| neuroblastoma-derived AMV related oncogene/MYCN | 164840 | NM_005378 | 2p24.1 |
| GRFS related oncogene/FGR | 164940 | NM_005417 | 1p36.2-p36.1 |
| v-abl oncogene homolog 2/ABL2 | 164690 | M14904 | 1q24-q25 |
| v-abl oncogene homolog 1/ABL1 | 189980 | U07563 | 9q34.1 |
| murine sarcoma 3611-derived oncogene 1/v-RAF/ARAF1 | 311010 | NM_001654 | Xp11.4-p11.2 |
| murine sarcoma 3611-derived oncogene 2/v-RAF/ARAF2 | 164710 | ****** | 7p11.4-cen |
| AVE E26 oncogene homolog 1/v-ETS/ETS1 | 164720 | NM_005239 | 11q23.3 |
| AVE E26 oncogene homolog 2/v-ETS/ETS2 | 164740 | M11922 | 21q22.3 |

| | | | |
|---|---|---|---|
| AVE E26 oncogene homolog related/ERG | 165080 | NM_004449 | 21q22.3 |
| AVE E26 oncogene homolog ETS related/ELK1 | 311040 | M25269 | Xp11.2 |
| SFFV virus-induced erythroleukemia oncogene/SP11 | 165170 | NM_003120 | 11p12-p11.22 |
| VAV oncogenem 1/VAV1 | 164875 | NM_005428 | 19p13.3-p13.2 |
| VAV oncogenem 2/VAV2 | 600428 | NM_003371 | 9q34 |
| ASV oncokene homolog/SRC | 190090 | NM_005417 | 20q12-q13 |
| YSV oncogene homolog 1/YES1 | 164880 | NM_005433 | 18p11.3 |
| ASV oncogene homolog/SKI | 164780 | NM_003036 | 1q22-q24 |
| ASV oncogene homolog-like/SKIL/SNO | 165340 | NM_005414 | ****** |
| lung carcinoma-derived AMV oncogene homolog 1/MYCL1 | 164850 | M19720 | 1p34.3 |
| AMV oncogene homolog 1-like/MYCL2 | 164865 | M64786 | 7p15 |
| MCF.2 cell line-derived transforming sequence/MCF2 | 311030 | J03639 | Xq27 |
| 1/THRA | 190120 | M24898 | 17q11.2 |
| 1/THRB | 190160 | S72623 | 3p24.3 |
| cancer Osaka thyroid (COT) oncogene/COT | 603259 | D14497 | 10p11.2 |
| CAS-BR-M murine ecotropic retroviral oncogene/CBL | 165360 | ****** | 11q23.3 |

| | | | | |
|---|---|---|---|---|
| | | teratoma oncogene TC21 | 600098 | M31470 | ****** |
| | | liver cancer oncogene/LCO | 165320 | ****** | 2q14-q21 |
| | | V-FES FSV/V-FPS FASV oncogene homolog/FES | 190030 | NM_002005 | 15q26.1 |
| | | glioma-associated oncogene homolog/GLI | 165220 | NM_005269 | 12q13.2-q13.3 |
| | | V-CRK ASV CT10 oncogene homolog/CRK | 164762 | NM_005206 | 17p13.3 |
| | | V-CRK ASV CT10 oncogene homolog-like/CRKL | 602007 | X59656 | 22q11 |
| | | epithelial cell transforming sequence 2 oncogene/ECT2 | 600586 | ****** | 3q26.1-q26.2 |
| | | V-RAF MSV oncogene homolog B1/BRAF | 164757 | AH003899 | 7q34 |
| | | S13 AEV oncogene homolog/SEA | 165110 | ****** | 11q13 |
| | | neuroblastoma suppressor/NBS | 256700 | NM_005380 | 1p36.3-p36.2 |
| | | hepatocyte growth factor receptor/oncogene MET/MET | 164860 | NM_000245 | 7q31 |
| | | MOS oncogene homolog/MOS | 190060 | NM_005372 | 8q11 |
| | | nuclear receptor coactivator/AIB1 | 601937 | NM_006534 | 20q12 |
| **Signaling** | | signal transducer and activator of transcription/STAT1 | 600555 | M97935 | 2q32.2-q32.3 |
| | | lipocortin 1/annexin 1 | 151690 | V00546 | 9q11-q22 |
| | | lipocortin 2/annexin 2 | 151740 | D00017 | 15q21-q22 |
| | | lipocortin 3/annexin 3 | 106490 | NM_005139 | 4q21 |

| | | | |
|---|---|---|---|
| lipocortin 5/annexin 5 | 131230 | NM_001154 | 4q26-q28 |
| lipocortin 7/annexin 7 (splice variant 1) | 186360 | NM_004034 | 10q21.1-q21.2 |
| lipocortin 7/annexin 7 (splice variant 2) | 186360 | NM_001156 | 10q21.1-q21.2 |
| BCL2 | 151430 | M13994 | 18q21.3 |
| BCL-X/BCLX | 600039 | Z23115 | ****** |
| BCL2 associated protein/BAX | 600040 | L22473 | 19q13.3-q13.4 |
| BCL2-antagonist/killer 1/BAK1 | 600516 | NM_001188 | 6p21.3-p21.2 |
| BCL2-associated athanogene 1/BAG1 | 601497 | NM_004323 | 9p12 |
| BCL2-associated athanogene 2/BAG2 | 603882 | NM_004282 | ****** |
| BCL2-associated athanogene 3/BAG3 | 603883 | AF095193 | ****** |
| BCL2-associated athanogene 4/BAG4 | 603884 | AF095194 | ****** |
| BCL2-associated athanogene 5/BAG5 | 603885 | AF095195 | ****** |
| BCL-X/BCL-2 binding protein/BAD | 603167 | AF021792 | ****** |
| BCL2-like 1/BCL2L1 | 600039 | NM_001191 | ****** |
| BCL2-like 2/BCL2L2 | 601931 | NM_004050 | 14q11.2-q12 |
| BCL2-like 11 (apoptosis facilitator)/BCL2L11 | 603827 | NM_006538 | ****** |
| BCL2-related protein A1/BCL2A1 | 601056 | Y09397 | 15q24.3 |
| BCL2-interacting protein harikari/HRK | 603447 | NM_003806 | ****** |
| Bcl-2 interacting killer/BIK | 603392 | U34584 | ****** |

| | | | |
|---|---|---|---|
| apoptosis inhibitor 1/API1 | 601712 | NM_001166 | 11q22-q23 |
| apoptosis inhibitor 2/API2 | 601721 | NM_001165 | 11q22-q23 |
| apoptosis inhibitor 3/API3 | 300079 | NM_001167 | Xq25 |
| apoptosis inhibitor 4/API4 | 603352 | NM_001168 | ****** |
| secreted apoptosis related protein 1/SARP1/SFRP2 | 604157 | AF017986 | 4q31.3 |
| secreted apoptosis related protein 2/SARP2/SFRP1 | 604156 | AF017987 | 8p12-p11.1 |
| secreted apoptosis related protein 3/SARP3/SFRP5 | 604158 | AF017988 | 10q24.1 |
| programmed cell death 1/PDCD1 | 600244 | NM_005018 | 2q37.3 |
| programmed cell death 2/PDCD2 | 600866 | NM_002598 | 6q27 |
| programmed cell death 8/apoptosis-inducing factor/AIF/PDCD8 | 300169 | NM_004208 | Xq25-q26 |
| map-kinase activating death domain/DENN | 603584 | U44953 | 11p11.2 |
| death-associated protein/DAP | 600954 | NM_004394 | 5p15.2 |
| death-associated protein kinase 1/DAPK1 | 600831 | NM_004938 | 9q34.1 |
| death associated protein 3/DAP3 | 602074 | NM_004632 | 1q21 |
| death-associated protein kinase 3/DAPK3 | 603289 | NM_001348 | 19q13.3 |

| | | | |
|---|---|---|---|
| death associated protein 6/DAXX/DAP6 | 603186 | AF050179 | 6p21.3 |
| defender against cell death 1/DAD1 | 600243 | NM_001344 | 14q11-q12 |
| BH3 interacting domain death agonist/BID | 601997 | NM_001196 | 22q11.2 |
| TNF receptor-1 associated protein/TRADD | 603500 | L41690 | 16q22 |
| CASP2 and RIPK1 domain containing adaptor with death domain/CRADD | 603454 | NM_003805 | 12q21.33-q23.1 |
| neuronal apoptosis inhibitory protein/NAIP | 600355 | NM_004536 | 5q12.2-q13.3 |
| RING finger protein/ROC1 | 603814 | AF142059 | ****** |
| RING finger protein/ROC2 | 603863 | AF142060 | 3q22-q24 |
| tumor protein p53/TP53 | 191170 | X02469 | 17p13.1 |
| apoptosis linked gene/calcium binding protein/ALG2 | 601057 | AF035606 | 5p15.2-pter |
| requiem, apoptosis response zinc finger gene/REQ | 601671 | NM_006268 | 11q13 |
| Fas (TNFRSF6)-associated via death domain/FADD | 602457 | NM_003824 | 11q13.3 |
| chromosome segregation 1 (yeast homolog)-like/CSE1L | 601342 | NM_001316 | 20q13 |
| superfamily, member 6/FAS/TNFRSF6 | 134637 | NM_000043 | 10q24.1 |
| apoptotic protease activating factor/APAF1 | 602233 | NM_001160 | ****** |
| receptor-interacting serine-threonine kinase 1/RIPK1 | 603453 | NM_003804 | ****** |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | receptor-interacting serine-threonine kinase 2/RIPK2 | 603455 | NM_003821 | 8q21 |
| | | | apoptosis response protein/PAWR | 601936 | NM_002583 | 12q21 |
| | | | apoptosis-related cystein protease 1/caspase 1/CASP1 | 147678 | L27475 | 11q22.2-q22.3 |
| | | | apoptosis-related cystein protease 1/caspase 1/CASP2 | 600639 | ****** | 7q35 |
| | | | apoptosis-related cystein protease 1/caspase 1/CASP3 | 600636 | NM_004346 | 4q35, 4q33-q35.1 |
| | | | apoptosis-related cystein protease 1/caspase 1/CASP4 | 602664 | NM_004347 | 11q22.2-q22.3 |
| | | | apoptosis-related cystein protease 1/caspase 1/CASP5 | 602665 | NM_004347 | 11q22.2-q22.3 |
| | | | apoptosis-related cystein protease 1/caspase 1/CASP6 | 601532 | NM_001226 | 4q25-q25 |
| | | | apoptosis-related cystein protease 1/caspase 1/CASP7 | 601761 | NM_001227 | 10q25.1-q25.2 |
| | | | apoptosis-related cystein protease 1/caspase 1/CASP8 | 601763 | NM_001228 | 2q33-q34 |
| | | | apoptosis-related cystein protease 1/caspase 1/CASP9 | 602234 | ****** | ****** |
| | | | apoptosis-related cystein protease 1/caspase 1/CASP10 | 601762 | NM_001230 | 2q33-q34 |
| | | | apoptosis-related cystein protease 1/caspase 1/CASP13 | 603653 | NM_003723 | ****** |

**Table 2. Central Nervous System Gene List**

| Class | Pathway | Function | Name | OMIM | GID | Locus |
|---|---|---|---|---|---|---|

| | | | | | | |
|---|---|---|---|---|---|---|
| | | Concentration of Transmitter in Vesicles | vacuolar ATPase subunit 1 | 300197 | NM_001183 | ****** |
| | | | vacuolar ATPase subunit H | 603931 | NM_003945 | ****** |
| | | | vacuolar ATPase subunit D | 603097 | NM_004691 | ****** |
| | | | vacuolar ATPase subunit C | 108745 | NM_001694 | 16p13.3 |
| | | | vacuolar ATPase subunit F | 603717 | ****** | ****** |
| | | | vacuolar ATPase subunit E | 108746 | NM_001696 | 22q11.2 |
| | | | vacuolar ATPase subunit B | 192132 | AH007312 | 2cen-q13 |
| | | | vacuolar ATPase subunit N | 192130 | NM_001991 | 17q21 |
| | | | chromogranin A | 118910 | NM_001275 | 14q32 |
| | | | chromogranin B | 118920 | NM_001819 | 20pter-p12 |
| | | | chromogranin C/secretogranin 2 | 118930 | ****** | 2q35-q36 |
| | | | carboxypeptidase E/CPE | 114855 | NM_001873 | Chr.4 |
| | | | secretory granule neuroendocrine protein 1/SGNE1 | 173120 | NM_003020 | 15q11-q15 |
| | | | Nerve growth factor inducible protein/VGF | 602186 | NM_003378 | 7q22 |
| | | | neuronal calcium sensor 1/NCS1/frequenin | 603315 | ****** | ****** |
| | | | amphiphysin | 600418 | NM_001635 | 7p14-p13 |
| | | | synapsin 1 | 313440 | ****** | Xp11.4-p11.2 |
| | | | synapsin 2 | 600755 | ****** | 3p |
| | | | synapsin 3 | 602705 | NM_003490 | 22q12.3 |
| | | | syntaxin 1A | 186590 | D37932 | 7q11.2 |
| | | | syntaxin 1B | 601485 | ****** | 16p11.2 |
| | | | syntaxin 3A | 600876 | NM_004177 | ****** |
| | | | syntaxin 4A | 186591 | NM_004604 | ****** |
| | | | syntaxin 5A | 603189 | NM_003164 | ****** |
| | | | syntaxin 6 | 603944 | NM_005819 | ****** |
| | | | syntaxin 7 | 603217 | NM_003569 | Chr.6 |
| | | | syntaxin 10 | 603765 | NM_003765 | ****** |

| | | | | | |
|---|---|---|---|---|---|
| | **Packaging of Neurotransmitter into Vesicles and Release** *(common to all small molecule neurotransmitters excluding nitric oxide)* | **Neurotransmitter Release** | syntaxin 16 | 603666 | AF038897 | ****** |
| | | | syntaxin binding protein 1 | 602926 | NM_003165 | 9q34.1 |
| | | | syntaxin binding protein 2 | 601717 | U63533 | 9p13.3-p13.2 |
| | | | neurexin 1 | 600565 | AB011182 | ****** |
| | | | neurexin 2 | 600566 | ****** | ****** |
| | | | neurexin 3 | 600567 | NM_004796 | ****** |
| | | | synaptotagmin 1/SYT1 | 185605 | NM_005639 | 12cen-q21 |
| | | | synaptotagmin 2/SYT2 | 600104 | ****** | 1q |
| | | | synaptotagmin 3/SYT3 | 600327 | ****** | 19q |
| | | | synaptotagmin 4/SYT4 | 600103 | ****** | 5q |
| | | | synaptotagmin 5/SYT5 | 600782 | ****** | 11p |
| | | | synaptobrevin 1/VAMP1 | 185880 | AH002992 | 12p |
| | | | synaptobrevin 2/VAMP2 | 185881 | AH002993 | 17pter-p12 |
| | | | cellubrevin/VAMP3 | 603657 | ****** | ****** |
| | | | endobrevin/VAMP8 | 603177 | ****** | ****** |
| | | | N-ethylmaleimide sensitive | 601633 | ****** | 17q21-q22 |
| | | | soluble NSF-attachment protein gamma/gamma SNAP | 603216 | NM_003826 | ****** |
| | | | soluble NSF-attachment protein alpha/alpha SNAP | 603215 | NM_003827 | ****** |
| | | | synaptosomal-associated protein 23/SNAP23 | 602534 | NM_003825 | ****** |
| | | | synaptosomal-associated protein 25/SNAP25 | 600322 | NM_003081 | 20p11.2 |
| | | | Golgi SNARE 27/membrin | 604027 | NM_004287 | 17q21 |
| | | | Golgi SNARE 28 | 604026 | ****** | 17q11 |
| | | | secretion deficient 22C | 604028 | AF039568 | ****** |
| | | | secretion deficient 22L1 | 604029 | NM_004892 | 1q21.2-q21.3 |
| | | | bassoon homolog | 604020 | NM_003458 | 3p21.31 |

| | | | |
|---|---|---|---|
| voltage-dependent Ca channel L type subunit 1A | 601011 | NM_000068 | 19p13 |
| voltage-dependent Ca channel N type subunit 1B | 601012 | NM_000718 | 9q34 |
| voltage-dependent Ca channel L type subunit 1D | 114206 | NM_000720 | 3p14.3 |
| large conductance Ca-activated K channel M type subunit 1B | 603951 | NM_004137 | 5q34 |
| large conductance Ca-activated K channel M type subunit 1A | 600150 | U09384 | Chr.10 |
| RAS-associated protein RAB1 | 179508 | NM_004161 | 2p14-p13.4 |
| RAS-associated protein RAB2 | 179509 | M28213 | ****** |
| RAS-associated protein RAB3A | 179490 | NM_002866 | 19p13.1-p12 |
| RAS-associated protein RAB3B | 179510 | NM_002867 | 1p32-p31 |
| RAS-associated protein RAB4 | 179511 | NM_004578 | 1q42-q43 |
| RAS-associated protein RAB5A | 179512 | NM_004162 | 3p24-p22 |
| RAS-associated protein RAB6 | 179513 | NM_002869 | 2q14-q21 |
| agrin | 103320 | S44195 | 1pter-p32 |
| synaptic vesicle protein 2/SV2 | 185860 | ****** | ****** |
| synaptic vesicle protein 2B/SV2B | 185861 | ****** | ****** |
| axonal transporter of synaptic vesicles/ATSV | 601255 | NM_004321 | 2q37 |
| synaptophysin/SVP | 313475 | X06389 | p11.23-p11.2 |
| phosphatidylinositol-4-kinase alpha catalytic subunit/PIK4CA | 600286 | NM_002650 | ****** |
| phosphatidylinositol-4-kinase beta catalytic subunit/PIK4CB | 602758 | NM_002651 | 1q21.1-q21.3 |
| dynamin 1/DNM1 | 602377 | NM_004408 | 9q34 |
| paired basic amino acid cleaving enzyme/furin/PACE | 136950 | X04329 | 15q25-q26 |

| | | | | | |
|---|---|---|---|---|---|
| | | | beta 1 adaptin/ADTB1 | 600157 | NM_001127 | 22q12 |
| | | | beta 3A adaptin/ADTB3A | 603401 | NM_001284 | ****** |
| | | | gamma adaptin/ADTBG | 603533 | NM_001283 | 16q23 |
| | | | gamma 2 adaptin/ADPTG2 | 603534 | NM_001283 | NM_001283 |
| | | | human homolog of S. cerevisea VT11 | 603207 | AF035824 | ****** |
| | | | huntingtin (Huntington disease)/HD | 143100 | NM_002111 | 4p16.3 |
| | | | huntingtin-associated protein 1/HAP2 | 600947 | AF040723 | 17q |
| | | **General Signal Transduction** | phospholipase C beta 4/PLCB4 | 600810 | NM_000933 | 20p12 |
| | | | calmodulin 1/CALM1 | 114180 | AH005370 | 14q24-q31 |
| | | | calmodulin 2/CALM2 | 114182 | NM_001743 | 2p21.3-p21.1 |
| | | | calmodulin 3/CALM3 | 114183 | NM_005184 | 9q13.2-q13.3 |
| | | | calcium/calmodulin dependent protein kinase II alpha/CAMK2A | 114078 | ****** | ****** |
| | | | calcium/calmodulin dependent protein kinase II gamma/CAMK2G | 602123 | NM_001222 | 10q22 |
| | | | calcium/calmodulin dependent protein kinase IV/CAMK4 | 114080 | ****** | 5q21-q23 |
| | | **Biosynthesis** | Glutaminase | 138280 | AB020645 | 2q32-q34 |
| | | | Glutamate dehydrogenase 1 | 138130 | X07674 | 10q23.3 |
| | | | Glutamate dehydrogenase 2 | 300144 | U08997 | Xq25 |
| | | | Glutamate Receptor, Ionotropic, Ampa 1; Gria1 | 138248 | M64752 | 5q33 |
| | | | Glutamate Receptor, Ionotropic, Ampa 2; Gria2 | 138247 | L20814 | 4q32-q33 |
| | | | Glutamate Receptor, Ionotropic, Ampa 3; Gria3 | 305915 | X82068 | Xq25-q26 |
| | | | Glutamate Receptor, Ionotropic, Ampa 4; Gria4 | 138246 | NM_000829 | 11q22-q23 |

| | | | | | |
|---|---|---|---|---|---|
| | | | Glutamate Receptor, Ionotropic, Delta 2; Grid2 | 602368 | AF009014 | 4q22 |
| | | | Glutamate Receptor, Ionotropic, Kainate 1; Grik1 | 138245 | U16125 | 21q22 |
| | | | Glutamate Receptor, Ionotropic, Kainate 2; Grik2 | 138244 | S75105 | 6q21 |
| | | | Glutamate Receptor, Ionotropic, Kainate 3; Grik3 | 138243 | U16127 | 1p34-p33 |
| | | | Glutamate Receptor, Ionotropic, Kainate 4; Grik4 | 600282 | S67803 | 11q22.3 |
| | | | Glutamate Receptor, Ionotropic, Kainate 5; Grik5 | 600283 | S40369 | 19q13.2 |
| | | | Glutamate Receptor, Ionotropic, N-Methyl-D-Asp 1; Grin1 | 138249 | L13266 | 9q34.3 |
| | | | Glutamate Receptor, Ionotropic, N-Methyl-D-Asp 2a; Grin2a | 138253 | U09002 | 16p13 |
| | | | Glutamate Receptor, Ionotropic, N-Methyl-D-Asp 2b; Grin2b | 138252 | U28758 | 12p12 |
| | | | Glutamate Receptor, Ionotropic, N-Methyl-D-Asp 2c; Grin2c | 138254 | L76224 | 17q25 |
| | | | Glutamate Receptor, Ionotropic, N-Methyl-D-Asp 2d; Grin2d | 602717 | U77783 | 19q13.1-qter |
| | | | Glutamate Receptor, Ionotropic, N-Methyl-D-Asp A; Grina | 138251 | ****** | 8q24.3 |
| | | | Glutamate Receptor, Metabotropic 2/G protein-coupled/ Grm2 | 604099 | L35318 | ****** |
| | | | Glutamate Receptor, Metabotropic 3/G protein-coupled/Grm3 | 601115 | X77748 | 7q21.1-q21.2 |
| | | | Glutamate Receptor, Metabotropic 4/G protein-coupled/Grm4 | 604100 | X80818 | ****** |

**Glutamate (NMDA) Pathway**

**Receptors**

| | | | | | |
|---|---|---|---|---|---|
| | | | Glutamate Receptor, Metabotropic 5/G protein-coupled/Grm5 | 604102 | D28538 | ****** |
| | | | Glutamate Receptor, Metabotropic 6/G protein-coupled/Grm6 | 604096 | U82083 | ****** |
| | | | Glutamate Receptor, Metabotropic 7/G protein-coupled/Grm7 | 604101 | X94552 | ****** |
| | | | Glutamate Receptor, Metabotropic 8/G protein-coupled/Grm8 | 601116 | U95025 | 7q31.3-q32.1 |
| | | **Reuptake** | Solute Carrier Family 1, Member 1; Slc1a1 | 133550 | U08989 | 9p24 |
| | | | Solute Carrier Family 1, Member 2; Slc1a2 | 600300 | U03505 | 11p13-p12 |
| | | | Solute Carrier Family 1, Member 3; Slc1a3 | 600111 | U03504 | 5p13 |
| | | **Catabolism** | Glutamine Synthetase | 138290 | X59834 | 1q31 |
| | | | soluble glutamate oxaloacetate transaminase 1/GOT1 | 138180 | NM_002079 | 0q24.1-q25.1 |
| | | | mitochondrial glutamate oxaloacetate transaminase 2/GOT2 | 138150 | NM_002080 | 16q21 |
| | | **Biosynthesis** | aromatic L-Amino Acid Decarboxylase/AADC | 107930 | M76180 | 7p11 |
| | | | tryptophan hydroxylase/TPH | 191060 | X52836 | 11p15.3-p14 |
| | | | 14-3-3 protein ETA | 113508 | X78138 | 22q12 |
| | | | 14-3-3 protein ZETA | 601288 | M86400 | 2p25.2-p25.1 |
| | | | 14-3-3 protein BETA | 601289 | X57346 | 20q13.1 |
| | | | 14-3-3 protein SIGMA | 601290 | X57348 | ****** |
| | | | serotonin 5-HT receptors 5-HT1A, G protein-coupled | 109760 | X57829 | 5q11.2-q13 |

| | | | | | |
|---|---|---|---|---|---|
| Serotonin | Receptors | serotonin 5-HT receptors 5-HT1B, G protein-coupled | 182131 | M81590 | 6q13 |
| | | serotonin 5-HT receptors 5-HT1C, G protein-coupled | 312861 | U49516 | Xq24 |
| | | serotonin 5-HT receptors 5-HT1D, G protein-coupled | 182133 | M81590 | 1p36.3-p34.3 |
| | | serotonin 5-HT receptors 5-HT1E, G protein-coupled | 182132 | M91467 | 6q14-q15 |
| | | serotonin 5-HT receptors 5-HT1F, G protein-coupled | 182134 | L05597 | 3p12 |
| | | serotonin 5-HT receptors 5-HT2A, G protein-coupled | 182135 | D87030 | 13q14-q21 |
| | | serotonin 5-HT receptors 5-HT2B, G protein-coupled | 601122 | X77307 | 2q36.3-q37.1 |
| | | serotonin 5-HT receptors 5-HT2C, G protein-coupled | 312861 | U49516 | Xq24 |
| | | serotonin 5-HT receptors 5-HT3, gated ion channel | 182139 | D49394 | 1q23.1-q23.2 |
| | | serotonin 5-HT receptors 5-HT4, G protein-coupled | 602164 | Y08756 | 5q31-q33 |
| | | serotonin 5-HT receptors 5-HT5a, G protein-coupled | 601305 | X81411 | 7q36.1 |
| | | serotonin 5-HT receptors 5-HT6, G protein-coupled | 601109 | L41147 | 1p36-p35 |
| | | serotonin 5-HT receptors 5-HT7, G protein-coupled | 182137 | L21195 | 10q21-q24 |
| | Reuptake | serotonin transporter | 182138 | X70697 | 17q11.1-q12 |
| | Catabolism | monoamine oxidase A; MAOA | 309850 | M69226 | Xp11.23 |
| | | monoamine oxidase B; MAOB | 309860 | M69177 | Xp11.23 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | Catabolism | serotonin N-Acetyltransferase/SNAT | 600950 | U40347 | 17q25 |
| | | | tryptophan 2,3-dioxygenase/TDO2 | 191070 | NM_005651 | 4q31-q32 |
| Small Molecule Neurotransmitters | Dopamine Pathway | Biosynthesis | Aromatic L-Amino Acid Decarboxylase/AADC/dopa decarboxylase | 107930 | M76180 | 7p11 |
| | | | Tyrosine Hydroxylase | 191290 | X05290 | 11p15.5 |
| | | Receptors | Dopamine Receptor D1 | 126449 | X58987 | 5q35.1 |
| | | | Dopamine Receptor D2/DRD2 | 126450 | NM_000795 | 11q23 |
| | | | Dopamine Receptor D3/DRD3 | 126451 | U32499 | 3q13.3 |
| | | | Dopamine Receptor D4 | 126452 | L12398 | 11p15.5 |
| | | | Dopamine Receptor D5 | 126453 | M67439 | 4p16.1-p15.3 |
| | | Reuptake | Dopamine Transporter/ DAT1 | 126455 | L24178 | 5p15.3 |
| | | Catabolism | Dopamine Beta-Hydroxylase/monooxygenase | 223360 | Y00096 | 9q34 |
| | | | Catechol-O-Methyltransferase | 116790 | M58525 | 22q11.2 |
| | | | Monoamine Oxidases A | 309850 | M69226 | Xp11.23 |
| | | | Monoamine Oxidases B | 309860 | M69177 | Xp11.23 |
| | | | Phenol Sulfotransferase 1 | 171150 | L10819 | 6p12.1-p11.2 |
| | | | Phenol Sulfotransferase 2 | 601292 | X78282 | 16p12.1-p11.2 |
| | | | Phenol Sulfotransferase 3 | 600641 | L19956 | 16p11.2 |
| | | Biosynthesis | dopamine beta hydroxylase | 223360 | Y00096 | 9q34 |
| | | | phenylethanolamine-N- | 171190 | NM_002686 | 17q21-q22 |
| | | | tyrosine Hydroxylase | 191290 | X05290 | 11p15.5 |
| | | | alpha-1a-adrenergic receptor; | 104219 | M76446 | Chr.20 |
| | | | alpha-1b-adrenergic receptor; | 104220 | L31773 | 5q33 |
| | | | alpha-1c-adrenergic receptor; | 104221 | D25235 | 8p21 |
| | | | alpha-1d-adrenergic receptor; | 104222 | M76446 | 20p13 |
| | | | alpha-2a-adrenergic receptor; | 104210 | M18415 | 10q24-q26 |
| | | | alpha-2b-adrenergic receptor; | 104260 | AF005900 | Chr.2 |

| | | | | | |
|---|---|---|---|---|---|
| Epinephrine and Norepinephrine Pathway | Receptors | alpha-2c-adrenergic receptor; | 104250 | J03853 | 4q16.1 |
| | | beta-1-adrenergic receptor; Adrb1 | 109630 | J03019 | 10q24-q26 |
| | | Beta-2-Adrenergic Receptor; Adrb2 | 109690 | M15169 | 5q32-q34 |
| | | beta-adrinergic receptor kinase 1/BARK1 | 109635 | NM_001619 | 11cen-q13 |
| | | Beta-2-Adrenergic Receptor-Like Protein G-21 | 109760 | X57829 | 5q11.2-q13 |
| | | Beta-3-Adrenergic Receptor; Adrb3 | 109691 | X70811 | 8p12-p11.2 |
| | | Beta-Adrenergic Receptor Kinase 1; Adrbk1 | 109635 | X61157 | 11cen-q13 |
| | | Beta-Adrenergic Receptor Kinase 2; Adrbk2 | 109636 | X69117 | 22q11 |
| | Reuptake | Vesicular Amine Transporter 2; VAT2 | 193001 | L09118 | 10q25 |
| | | Vesicular Amine Transporter 1; VAT1 | 193002 | ****** | 8p21.3 |
| | | Solute carrier family 6 , member 5/SLC6A2/NAT1/NET1 | 163970 | NM_001043 | 16q12.2 |
| | Catabolism | Monoamine Oxidase A; MAOA | 309850 | M69226 | Xp11.23 |
| | | Monoamine Oxidase B; MAOB | 309860 | M69177 | Xp11.23 |
| | | Catechol-O-Methyltransferase | 116790 | M58525 | 22q11.2 |
| | Biosynthesis | Choline acetyltransferase/CHAT | 118490 | NM_003055 | 10q11.2 |
| | | carnitine acetyltransferase/CRAT | 600184 | NM_004003 | 9q34.1 |
| | | apolipoprotein E | 107741 | NM_000041 | 19q13.2 |
| | | Cholinergic Receptor, Muscarinic, 1; CHRM1 | 118510 | X15263 | 11q13 |
| | | Cholinergic Receptor, Muscarinic, 2; CHRM2 | 118493 | U19800 | 7q35-q36 |
| | | Cholinergic Receptor, Muscarinic, 3; CHRM3 | 118494 | U29589 | 1q41-q44 |

| | | | | | |
|---|---|---|---|---|---|
| Acetylcholine Pathway | Receptors | Cholinergic Receptor, Muscarinic, 4; CHRM4 | 118495 | M16405 | 11p12-p11.2 |
| | | Cholinergic Receptor, Muscarinic, 5; CHRM5 | 118496 | AF026263 | 15q26 |
| | | Nicotinic, Cholinergic receptor alpha 1 | 100690 | X70108 | 2q24-q32 |
| | | Nicotinic, Cholinergic receptor alpha 2 | 118502 | U62431 | Chr.8 |
| | | Nicotinic, Cholinergic receptor alpha 3 | 118503 | X53559 | 15q24 |
| | | Nicotinic, Cholinergic receptor alpha 4 | 118504 | U62433 | 20q13.2-q13.3 |
| | | Nicotinic, Cholinergic receptor alpha 5 | 118505 | M83712 | 15q24 |
| | | Nicotinic, Cholinergic receptor alpha 7/CHRNA7 | 118511 | U40583 | 15q14 |
| | | Nicotinic, Cholinergic receptor beta 1 | 100710 | X14830 | 17p12-p11 |
| | | Nicotinic, Cholinergic receptor beta 2 | 118507 | Y08415 | 1p21 |
| | | Nicotinic, Cholinergic receptor beta 3 | 118508 | X67513 | 8p11.2 |
| | | Nicotinic, Cholinergic receptor beta 4 | 118509 | X68275 | 15q24 |
| | | Nicotinic, Cholinergic receptor epsilon polypeptide | 100725 | X66403 | Chr.17 |
| | | Nicotinic, Cholinergic receptor, | 100720 | X55019 | 2q33-q34 |
| | | Nicotinic, Cholinergic receptor, | 100730 | NM_005199 | 2q33-q34 |
| | Reuptake | Vesicular acetylcholine transporter | 600336 | NM_003055 | 10q11.2 |
| | Catabolism | Acetylcholinesterase/ACHE | 100740 | M55040 | 7q22 |
| | | butyrylcholinesterase 1/serum cholinesterase 1/BCHE1 | 177400 | NM_000055 | 3q26.1-q26.2 |
| | | butyrylcholinestarase 2/serum cholinesterase 2/BCHE2 | 177500 | ****** | 2q33-q35 |
| | Biosynthesis | Histidine Decarboxylase | 142704 | M60445 | 15q21-q22 |
| | Receptors | histamine H1 receptor/HRH1 | 600167 | NM_000861 | 3p21-p14 |
| | | histamine H2 receptor/HRH2 | 142703 | AB023486 | ****** |

| | | | | | |
|---|---|---|---|---|---|
| **Histaminergic Pathway** | | histamine H3 receptor/HRH3 | ****** | NM_007232 | ****** |
| | **Catabolism** | Histamine N- | ****** | NM_006895 | chr. 2 |
| | | Amine oxidase (copper-containing) 2/AOC2 | 602268 | D88213 | 17q21 |
| | | Amine oxidase (copper-containing) 3/AOC3 | 603735 | AF054985 | 17q21 |
| **Adenosine Pathway** | **Biosynthesis** | adenylosuccinate lyase/ADSL | 103050 | NM_000026 | 22q13.1 |
| | | adenylosuccinate synthetase/ADSS | 103060 | NM_001126 | 1cen-q12 |
| | **Receptors** | Adenosine A1 Receptor; Adora1/G protein-coupled | 102775 | L22214 | 1q32.1 |
| | | Adenosine A2 Receptor; Adora2a/G protein-coupled | 102776 | X68486 | 22q11.2 |
| | | Adenosine A2b Receptor; Adora2b/G protein-coupled | 600446 | X68487 | 17p12-p11.2 |
| | | Adenosine A3 Receptor; Adora3/G protein-coupled | 600445 | L20463 | 1p21-p13 |
| | | Adenosine A2 Receptor-like/ADORA2L1 | 102777 | ****** | 10q25.3-q26.3 |
| | | Purinergic Receptor P2x, Ligand-Gated Ion Channel, 1; P2rx1 | 600845 | NM_002558 | ****** |
| | | Purinergic Receptor P2x, Ligand-Gated Ion Channel, 3; P2rx3 | 600843 | Y07683 | 11q12 |
| | | Purinergic Receptor P2x, Ligand-Gated Ion Channel, 4; P2rx4 | 600846 | AF000234 | 12q24.32 |
| | | Purinergic Receptor P2x, Ligand-Gated Ion Channel, 5; P2rx5 | 602836 | NM_002561 | ****** |
| | | Purinergic Receptor P2x, Ligand-Gated Ion Channel, 7; P2rx7 | 602566 | Y09561 | 12q24 |

| | | | | | |
|---|---|---|---|---|---|
| | | | P2Y11 purinoceptor/G protein- | 602697 | ****** | ****** |
| | | | P2Y7 purinoceptor/leukotriene B4 receptor/G protein-coupled | 601531 | NM_000752 | 14q11.2-q12 |
| | | | P2Y2 purinoceptor/G protein-coupled | 600041 | U07225 | 1q13.5-q14.1 |
| | | | P2Y1 purinoceptor/G protein-coupled | 601167 | U42029 | 3q25 |
| | | | P2Y4 pyrimidinergic receptor/G protein-coupled | 300038 | NM_002565 | Xq13 |
| | | | P2Y6 pyrimidinergic receptor/purinoreceptor P2Y6/G protein-coupled/P2RY6 | 602451 | NM_004154 | 11q13.5 |
| | | **Reuptake** | Solute carrier family 29 (nucleosides), member 1/SLC29A1/ENT1 | 602193 | NM_004955 | 6p21.2-p21.1 |
| | | | Solute carrier family 29 (nucleosides), member 2/SLC29A2/ENT2 | 602110 | X86681 | 11q13 |
| | | **Catabolism** | adenosine deaminase | 102700 | NM_000022 | 20q13.11 |
| | | **Biosynthesis** | Glutamate decarboxylase 1 (brain, 67kD) | 266100 | M81883 | 2q31 |
| | | | Glutamate decarboxylase 2 (brain, 65kD) | 138275 | X69936 | 10p11.23 |
| | | | Glutamate decarboxylase 3 | 138276 | 138276 | 22q13 |
| | | | Gamma-Aminobutyric Acid Receptor, Beta-3; Gabrb3 | 137192 | M82919 | 15q11.2-q12 |
| | | | Gamma-Aminobutyric Acid Receptor, Alpha-3; Gabra3 | 305660 | S62908 | Xq28 |
| | | | Gamma-Aminobutyric Acid Receptor, Alpha-5; Gabra5 | 137142 | L08485 | 15q11.2-q12 |
| | | | Gamma-Aminobutyric Acid Receptor, Alpha-1; Gabra1 | 137160 | X14766 | 5q34-q35 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | **Gamma-Aminobutyric Acid Pathway** | **Receptors** | Gamma-Aminobutyric Acid Receptor, Alpha-6; Gabra6 | 137143 | S81944 | 5q31.1-q35 |
| | | | Gamma-Aminobutyric Acid B Receptor 1; Gabbr1 | 603540 | Y11044 | ****** |
| | | | Gamma-Aminobutyric Acid Receptor, Alpha-2; Gabra2 | 137140 | S62907 | 4p13-p12 |
| | | | Gamma-Aminobutyric Acid Receptor, Gamma-3; Gabrg3 | 600233 | NM_000816 | 15q11.2-q12 |
| | | | Gamma-Aminobutyric Acid Receptor, Beta-1; Gabrb1 | 137190 | X14767 | 4p13-p12 |
| | | | Gamma-Aminobutyric Acid Receptor, Pi; Gabrp | 602729 | U95367 | ****** |
| | | | Gamma-Aminobutyric Acid Receptor, Epsilon; Gabre | 300093 | Y09765 | Xq28 |
| | | | Gamma-Aminobutyric Acid Receptor, Alpha-4; Gabra4 | 137141 | U30461 | 4p14-q12 |
| | | | Gamma-Aminobutyric Acid Receptor, Beta-2; Gabrb2 | 600232 | S77553 | 5q34-q35 |
| | | | Gamma-Aminobutyric Acid Receptor, Gamma-2; Gabrg2 | 137164 | X15376 | 5q31.1-q33.1 |
| | | | Gamma-Aminobutyric Acid Receptor, Gamma-1; Gabrg1 | 137166 | ****** | 4p14-q21.1 |
| | | | Gamma-Aminobutyric Acid Receptor, Delta; Gabrd | 137163 | AF016917 | 1p |
| | | | Gamma-Aminobutyric Acid Receptor Subunit Rho1 | 137161 | M62400 | 6q14-q21 |
| | | | Gamma-Aminobutyric Acid Receptor Subunit Rho2 | 137162 | M86868 | 6q14-q21 |
| | | | Benzodiazepine receptor, peripheral | 109610 | NM_000714 | 22q13.31 |
| | | | diazepam binding inhibitor/DBI | 125950 | M15887 | 2q12-q21 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | **Reuptake** | Solute carrier family 6 ( GABA), member 1/SLC6A1 | 137165 | X54673 | 3p25-p24 |
| | | | Solute carrier family 1, member 6 (GABA/GLU)/SLC1A6 | 600637 | NM_005071 | ****** |
| | | | Solute carrier family 6 (betaine/GABA), member 12 | 603080 | U27699 | 12p13 |
| | | **Catabolism** | GABA-glutamate transaminase | 137150 | NM_000663 | ****** |
| | | | succinic semialdehyde dehydrogenase/SSADH | 271980 | NM_001080 | 6p22 |
| | **Glycine Pathway** | **Biosynthesis** | Sacrosine dehydrogenase | 268900 | ****** | 9q33-q34 |
| | | | Alanine-glyoxylate aminotransferase. | 259900 | NM_000030 | 2q36-q37 |
| | | | cytosolic serine hydroxymethyltransferase 1/SHMT1 | 182144 | NM_004169 | 17p11.2 |
| | | | mitochondrial serine hydroxymethyltransferase 2/SHMT2 | 138450 | ****** | 12q13 |
| | | **Receptors** | Glycine Receptor, Alpha-1 Subunit; Glra1 | 138491 | X52009 | 5q32 |
| | | | Glycine Receptor, Alpha-2 Subunit; Glra2 | 305990 | X52008 | Xp22.1-p21.2 |
| | | | Glycine Receptor, Alpha-3 Subunit; Glra3 | 600421 | AF018157 | 4q33-q34 |
| | | | Glycine Receptor, Beta Subunit; Glrb | 138492 | U33267 | 4q31.3 |
| | | **Reuptake** | Solute carrier family 6, Member 9; SLC6A9 (glycine) | 601019 | S70612 | 1p33 |
| | | | Solute carrier family 6, Member 5; SLC6A5 (glycine) | 604159 | NM_004211 | ****** |
| | | | Glycine aminotransferase/glycine cleavage T protein/GAT | 238310 | NM_000481 | 3p21.2-p21.1 |

| | | | | | |
|---|---|---|---|---|---|
| | Catabolism | Glycine dehydrogenase/glycine cleavage P protein | 238300 | M63635 | 9p22 |
| | | Aminomethyl carrier/glycine cleavage H protein | 238330 | NM_004483 | ****** |
| | | Dihydrolipoamide dehydrogenase/glycine cleavage L | 238331 | ****** | ****** |
| Taurine | Biosynthesis | cysteine dioxygenase, type I/CDO1 | 603943 | NM_001801 | 5q22-q23 |
| | | sulfite oxidase/SUOX | 272300 | NM_000456 | ****** |
| | Receptors | solute carrier family 6, member 6/taurine transporter/SLC6A6 | 186854 | U16120 | 3p25-q24 |
| Melatonin | Biosynthesis | serotonin N-Acetyltransferase/SNAT | 600950 | U40347 | 17q25 |
| | | X-chromosomal acetylserotonin N-methyltransferase/ASMT | 300015 | NM_004043 | Xpter-p22.32 |
| | | Y-chromosomal acetylserotonin N-methyltransferase/ASMT | 402500 | ****** | Ypter-p11.2 |
| | | acetylserotonin N-methyltransferase-like/ASMTL | 300162 | NM_004192 | Xpter-p22.32 |
| | Receptors | melatonin receptor 1A/MTNR1A | 600665 | NM_005958 | 4q35.1 |
| | | melatonin receptor 1B/MTNR1B | 600804 | NM_005959 | 11q21-q22 |
| | Catabolism | tryptophan 2,3-dioxygenase/TDO2 | 191070 | NM_005651 | 4q31-q32 |
| Nitric Oxide Pathway | Biosynthesis | nitric oxide synthetase 1/NOS1 | 163731 | AH001515 | 2q24.2-q24.3 |
| | | nitric oxide synthetase 2A/NOS2A | 163730 | X85766 | 17cen-q11.2 |
| | | macrophage nitric oxide synthetase 2B/NOS2B | 600719 | AH006623 | 17p13.1-q25 |
| | | macrophage nitric oxide synthetase 2C/NOS2C | 600720 | 600720 | 17p13.1-q25 |
| | | nitric oxide synthetase 3/NOS3 | 163729 | AH001515 | 7q36 |
| | | chondrocyte nitric oxide synthetase 3/NOS4 | 163728 | X73029 | ****** |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | arginase/ARG1 | 207800 | NM_000045 | ****** |
| | | | arginase/ARG2 | 107830 | NM_001172 | 4q24.1-q24.3 |
| | General Metabolism for Peptide Hormones | Biosynthesis and Catabolism | membrane metalloendopeptidase/MME/neutral endopeptidase | 120520 | AH002677 | 3q21-q27 |
| | | | calpain, large polypeptide L3/CAPN3 | 114240 | NM_000070 | 5q15.1-q21.1 |
| | | | Leucyl/cystinyl aminopeptidase | 151300 | U62768 | ****** |
| | | | carboxypeptidase N polypeptide 1/CPN1 | 603103 | NM_001308 | chr. 10 |
| | | | carboxypeptidase N polypeptide 2/regulatory subunit/CPN2 | 603104 | J05158 | 8p23-p22 |
| | | | meprin alpha subunit/MEP1A | 600388 | NM_005925 | 6p21.2-p21.1 |
| | | | meprin beta subunit/MEP1B | 600389 | NM_005925 | 8q12.2-q12.3 |
| | | | prolyl endopeptidase/PREP | 600400 | NM_002726 | 6q22 |
| | | | neuroendocrine convertase 1/NEC1 | 162150 | D73407 | 5q14-q21 |
| | | | peptidylglycine alpha-amidating monooxygenase /PAM/NEC2 | 170270 | NM_000919 | 5q14-q21 |
| | | | paired basic amino acid cleaving enzyme/PACE/FUR | 136950 | X04329 | 15q25-q26 |
| | Opioids | Biosynthesis | proopiomelanocortin | 176830 | NM_000939 | 2p23.3 |
| | | | prepronociceptin/nociceptin/nosistatin/PNOC | 601459 | ****** | 8p21 |
| | | | preproenkephalin B/prodynorphin/PDYN | 131340 | NM_006211 | 20pter-p12.21 |
| | | | preproenkephalin A/proenkephalin/PENK | 131330 | NM_006211 | 8q23-q24 |
| | | | Opioid Receptor, Mu-1; Oprm1 | 600018 | NM_000914 | 6q24-q25 |
| | | | Opioid Receptor, Kappa-1; Oprk1 | 165196 | U17298 | 8q11.2 |



WO 00/50639

PCT/US00/01392

| | | | | | |
|---|---|---|---|---|---|
| | Receptors | opioid receptor-like 1/OPRL 1 | 602548 | X77130 | ****** |
| | | Opioid Receptor, Delta-1; Oprd1 | 165195 | U10504 | 1p36.1-p34.3 |
| | | Opioid Receptor, Sigma 1 | 601978 | U75283 | ****** |
| | | opioid binding cell adhesion molecule/OBCAM | 600632 | ****** | Chr.11 |
| | | G protein-coupled receptor 7/GPR7 | 600730 | U22491 | 0q11.2-q21.1 |
| | | G protein-coupled receptor 8/GPR8 | 600731 | U22492 | 20q13.3 |
| Oxytocin | Biosynthesis | Oxytocin | 167050 | M25650 | 20p13 |
| | Receptors | Oxytocin receptor | 167055 | X64878 | 3p26.2 |
| | Catabolism | Leucyl/cystinyl aminopeptidase | 151300 | U62768 | ****** |
| Cholecystokinin (CCK) | Biosynthesis | Cholecystokinin/CCK | 118440 | L00354 | 3pter-p21 |
| | Receptors | Cholecystokinin A receptor/CCKAR | 118444 | L13605 | 4p15.2-p15.1 |
| | | Cholecystokinin B receptor/CCKBR | 118445 | L08112 | 1p15.5-p15.4 |
| Neuropeptide Y (NPY) | Biosynthesis | Neuropeptide Y/NPY | 162640 | K01911 | 7p15.1 |
| | Receptors | Neuropeptide Y receptor Y1/NPY1R | 162641 | M84755 | 4q31.3-q32 |
| | | Neuropeptide Y receptor Y2/NPY2R | 162642 | U32500 | 4q31 |
| | | Neuropeptide Y receptor Y3/chemokine receptor 4/CXCR4 | 162643 | X71635 | 2q21 |
| | | Neuropeptide Y receptor Y5 | 602001 | U94320 | 4q31-q32 |
| | | Neuropeptide Y receptor Y6 | 601770 | D86519 | 5q31 |
| Leptin | Biosynthesis | leptin/LEP | 164160 | NM_000230 | 7q31.3 |
| | Receptor | leptin receptor/LEPR | 601007 | NM_002303 | 1p31 |
| Neurotensin Pathway | Biosynthesis | Neurotensin | 162650 | U91618 | 12q21 |
| | | prolyl endopeptidase/PREP | 600400 | NM_002726 | 6q22 |
| | Receptors | Neurotensin receptor | 162651 | X70070 | 20q13 |
| Tachykinin or Substance P or | Biosynthesis | Neurokinin A/Tachykinin 1 or 2/Substance P or K | 162320 | U37529 | 7q21-q22 |
| | | Neurokinin B/Tachykinin 3 | 162330 | ****** | 12q13-q21 |

| | | | | | |
|---|---|---|---|---|---|
| Neurokinin Pathway | Receptors | Tachykinin NK1 receptor/TACR1 | 162323 | M81797 | Chr.2 |
| | | Tachykinin NK2 receptor/TACR2 | 162321 | M57414 | 10q11-q21 |
| | | Tachykinin NK3 receptor/TACR3 | 162332 | M89473 | ****** |
| Bradykinin | Biosynthesis | kininogen/KNG | 228960 | ****** | 3q27 |
| | | kallikrein 1/KLK1 | 147910 | AH002853 | 9q13.2-q13.4 |
| | Receptor | bradykinin receptor B1/BDKRB1 G protein-coupled | 600337 | NM_000710 | 4q32.1-q32.2 |
| | | bradykinin receptor B2/BDKRB2 G protein-coupled | 113503 | NM_000623 | 4q32.1-q32.2 |
| Angiotensin | Biosynthesis | angiotensinogen | 106150 | NM_000029 | 1q42-q43 |
| | | renin/REN | 179820 | NM_000537 | 1q32 |
| | | renin-binding protein/RENBP | 312420 | D10711 | Xq28 |
| | | angiotensin converting enzyme/dipeptidyl carboxypeptidase | 106180 | NM_000789 | 17q23 |
| | Receptors | angiotensin receptor 1/AGTR1A | 106165 | M87290 | 3q21-q25 |
| | | angiotensin II receptor type 2/AGTR2 | 300034 | U10273 | Xq22-q23 |
| | | vascular angiotensin II receptor type 1B/AGTR1B | 600015 | NM_004835 | ****** |
| | Catabolism | prolylcarboxypeptidase/PRCP | 176785 | NM_005040 | 11q14 |
| Vasopressin | Biosynthesis | Arginine Vasopressin | 192340 | X03172 | 20p13 |
| | Receptors | Arginine Vasopressin Receptor 1A/AVPR1A | 600821 | AF030625 | 12q14-q15 |
| | | Arginine Vasopressin Receptor 1B/AVPR1B | 600264 | AF030512 | 1q32 |
| | | Arginine vasopressin receptor 2 | 304800 | AF030626 | Xq28 |
| | Catabolism | Leucyl/cystinyl aminopeptidase | 151300 | U62768 | ****** |
| | Biosynthesis | prepro-vasoactive intestinal | 192320 | AH003029 | 6q26-q27 |
| | | adenylate-cyclase activating polypeptide 1/ADCYAP1 | 102980 | NM_001117 | 18p11 |



WO 00/50639

PCT/US00/01392

| | | | | | | |
|---|---|---|---|---|---|---|
| Peptide Hormones | Vasoactive Intestinal Peptide | Receptor | vasoactive intestinal peptide receptor 1/VIPR1 | 192321 | ****** | 3p22 |
| | | | vasoactive intestinal peptide receptor 2/VIPR2 | 601970 | L40764 | 7q36.3 |
| | | | adenylate-cyclase activating polypeptide 1 receptor/ADCYAP1R1 | 102981 | D17516 | 7p14 |
| | Calcitonin and CGRP | Biosynthesis | calcitonin/calcitonin gene-related pepti | 114130 | M12667 | 1p15.2-p15.1 |
| | | | calcitonin/calcitonin gene-related pepti | 114160 | X02404 | 1p15.2-p15.2 |
| | | Receptor | calcitonin receptor/CALCR | 114131 | L00587 | 7q21.3 |
| | | | calcitonin receptor-like/CALCRL | 114190 | NM_005795 | ****** |
| | CRH | Biosynthesis | Corticotropin releasing hormone/CRH | 122560 | NM_000756 | 8q13 |
| | | | urocortin/UCN | 600945 | NM_003353 | Chr.2 |
| | | | urocortin 2/UCN2 | 604097 | AF104118 | ****** |
| | | Receptors | Corticotropin releasing hormone receptor 1 | 122561 | U16273 | 17q12-q22 |
| | | | Corticotropin releasing hormone receptor 2 | 602034 | NM_001883 | 7p21-p15 |
| | | | Corticotropin releasing hormone-binding protein | 122559 | X58022 | 5q11.2-q13.3 |
| | TRH | Biosynthesis | thyrotropin releasing hormone/TRH | 275120 | AH001523 | 3p |
| | | Receptors | thyrotropin releasing hormone receptor/G protein coupled/TRHR | 188545 | X75071 | 8q23 |
| | TSH | Biosynthesis | chorionic gonadotropin alpha chain/TSHA/CGA | 118850 | NM_000735 | 6q21.1-q23 |
| | | | thyroid stimulating hormone beta | 188540 | AH001548 | 1p13 |
| | | Receptor | thyroid stimulating hormone receptor | 603372 | NM_000369 | 14q31 |
| | | Biosynthesis | gonadotropin releasing hormone 1/LHRH/GNRH1 | 152760 | NM_000825 | 8p21-p11.2 |



WO 00/50639

PCT/US00/01392

| | | | | | |
|---|---|---|---|---|---|
| GRH | Biosynthesis | gonadotropin releasing hormone 2/LHRH/GNRH2 | 602352 | NM_001501 | 20p13 |
| | Receptor | gonadotropin releasing hormone receptor/G protein-coupled/LHRHR/GNRHR | 138850 | NM_000406 | 4q21.2 |
| FSH | Biosynthesis | follicle stimulating hormone-inhibin, beta A (activin A, activin AB alpha polypeptide)/INHBA | 136530 | AH002701 | 11p13 |
| | | activin A receptor, type I/ACVR1 | 147290 | NM_002192 | 7p15-p13 |
| | | activin A receptor, type IB/ACVR1B | 102576 | NM_001105 | 2q23-q24 |
| | | activin A receptor type II-like 1/ACVRL1 | 601300 | NM_004302 | 12q13 |
| | | activin typeII A receptor/ACVR2 | 601284 | NM_000020 | 12q11-q14 |
| | | activin A receptor, type IIB/ACVR2B | 102581 | D31770 | ****** |
| | | alpha-inhibin/INHA | 602730 | NM_001106 | 3p22-p21.3 |
| | | beta-B inhibin/beta C inhibin//INHBC | 147380 | M13144 | 2q33-q36 |
| | Receptor | follicle stimulating hormone receptor/FSHR | 601233 | M13437 | 12q13.1 |
| | | FSH primary response (LRPR1, rat) homolog 1/FSHPRH1 | 136435 | NM_000145 | 2p21-p16 |
| Somatostatin | Biosynthesis | Somatostatin | 300065 | NM_006733 | Xq22 |
| | | preprocortistatin | 182450 | J00306 | 3q28 |
| | Receptors | Somatostatin receptor 1/G protein-coupled | 602784 | NM_001302 | 1p36 |
| | | Somatostatin receptor 2 | 182451 | M81829 | 14q13 |
| | | Somatostatin receptor 3/adenyl cyclase coupled | 182452 | M81830 | 17q24 |
| | | Somatostatin receptor 4 | 182453 | M96738 | 22q13.1 |
| | | Somatostatin receptor 5 | 182454 | L07833 | 20p11.2 |
| | | | 182455 | D16827 | 16p13.3 |

| | | | | | |
|---|---|---|---|---|---|
| GHRH | Biosynthesis | growth hormone releasing hormone/GHRH | 139190 | AH002712 | 20q11.2 |
| | Receptor | growth hormone releasing hormone receptor/G protein-coupled/GHRHR | 139191 | U34195 | 7p15-p14 |
| Growth Hormone | Biosynthesis | growth hormone 1/somatotropin/GH1 | 139250 | NM_000515 | 17q22-q24 |
| | Receptor | growth hormone receptor/GHR | 600946 | NM_000163 | 5p13-p12 |
| ACTH | Biosynthesis | proopiomelanocortin | 176830 | NM_000939 | 2p23.3 |
| | Receptor | melanocortin 2 receptor/ACTH receptor/MC2R | 202200 | NM_000529 | 18p11.2 |
| Prolactin | Biosynthesis | prolactin/PRL | 176760 | NM_000948 | 6p22.2-p21.3 |
| | Receptor | prolactin receptor/PRLR | 176761 | NM_000949 | 5p13-p12 |
| Galanin | Biosynthesis | preprogalanin/GAL1 | 137035 | L11144 | 1q13.3-q13.3 |
| | Receptor | galanin receptor 1 (brain)/GALR1 | 600377 | NM_001480 | 18q23 |
| | | galanin receptor 2/GALR2 | 603691 | NM_003857 | 17q25.3 |
| | | galanin receptor 3 (brain)/GALR3 | 603692 | NM_003614 | 2q12.2-q13.1 |
| Bombesin | Biosynthesis | gastrin-releasing polypeptide/bombesin/GRP | 137260 | NM_002091 | 18q21 |
| | | neuromedin B/NMB | 162340 | M21551 | 15q22-qter |
| | Receptor | gastrin-releasing polypeptide receptor/G protein-coupled/GRPR | 305670 | D87058 | Xp22.3-p21.2 |
| | | neuromedin B receptor/g protein-coupled/NMBR | 162341 | ****** | 6q21-qter |
| | | bombesin-like receptor 3/BRS3 | 300107 | NM_001727 | Xq26-q28 |
| Glucagon Pathway | Biosynthesis | preproglucagon/GCG | 138030 | X03991 | 2q36-q37 |
| | Receptor | glucagon receptor/GCGR | 138033 | NM_000160 | 17q25 |
| | | glucagon-like peptide 1 | 138032 | U01156 | 6p21 |
| | | glucagon-like peptide 2 | 603659 | ****** | 17p13.3 |
| | Biosynthesis | carnosine synthase (6.3.2.11) | ****** | ****** | ****** |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Carnosine Pathway | Biosynthesis | homocarnosine synthase | ****** | ****** | ****** |
| | | Receptor | carnosine receptor | ****** | ****** | ****** |
| | | Catabolism | carnosinase/Xaa-his dipeptidase (3.4.13.3) | ****** | ****** | ****** |
| Steroid Hormones | Estrogen | Biosynthesis | cytochrome P450, subfamily XIX (androgen aromatase)/CYP19 | 107910 | NM_000103 | 15q21.1 |
| | | Receptors | estrogen receptor 1 (ESR1) | 133430 | M12674 | 6q25.1 |
| | | | estrogen receptor 2 (ESR2) | 601663 | X99101 | 14q |
| | | | estrogen-related receptor | 601998 | NM_004451 | 11q12 |
| | | | estrogen-related receptor beta/ESRRB | 602167 | NM_004452 | 14q24.3 |
| | | Catabolism | estrogen-preferring | 600043 | NM_005420 | 4q13.1 |
| | Testosterone /DHT | Biosynthesis | steroid 5-alpha-reductase 1/SRD5A1 | 184753 | AH003000 | 5p15 |
| | | | steroid 5-alpha-reductase 2/SRD5A2 | 264600 | NM_000348 | 2p23 |
| | | | aldo-keto reductase family 1, member C4/3-a hydroxysteroid dehydrogenase/AKR1C4 | 600451 | ****** | 10p15-p14 |
| | | Receptors | androgen receptor | 313700 | M20132 | Xq11-q12 |
| | | Catabolism | UDP glycosyltransferase 2 family, polypeptide B17/UGT2B17 | 601903 | NM_001077 | 4q13 |
| | Glucocorticosteroid | Biosynthesis | cytochrome P450, subfamily XXI (steroid 21-a-hydroxylase)/CYP21 | 201910 | M13936 | 6p21.3 |
| | | | cytochrome P450, subfamily XIB, polypeptide 2 (steroid 11-b-hydroxylase)/CYP11B2 | 124080 | NM_000498 | 8q21 |
| | | Receptors | glucocorticoid receptor/GRL | 138040 | NM_000176 | 5q31 |
| | | Metabolism | corticosteroid binding globulin precursor/CBG | 122500 | NM_001756 | 14q32.1 |
| | | | hydroxy-D-5-steroid dehydrogenase, 3 b- and steroid D-isomerase 2/HSD3B2 | 201810 | NM_000198 | 1p13.1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | voltage-dependent calcium channel, P/Q type, alpha 1A | 601011 | NM_000068 | 19p13 |
| | | | calcium channel, voltage-dependent, L type, alpha 1B subunit/CACNA1B | 601012 | NM_000718 | 9q34 |
| | | | calcium channel, voltage-dependent, L type, alpha 1C subunit/CACNA1C | 114205 | NM_000719 | 12p13.3 |
| | | | calcium channel, voltage-dependent, L type, alpha 1D subunit/CACNA1D | 114206 | NM_000720 | 3p14.3 |
| | | | L-type voltage dependent calcium channel alpha 1S subunit/CACNA1S | 114208 | NM_000069 | 1q32 |
| | | | calcium channel, voltage-dependent, beta 1 subunit/CACNB1 | 114207 | NM_000723 | 17q21-q22 |
| | | | voltage dependent calcium channel beta 2 subunit/CACNB2 | 600003 | U07139 | 10p12 |
| | | | voltage dependent calcium channel beta 3 subunit/CACNB3 | 601958 | NM_000725 | 12q13 |
| | | **Calcium Channels** | voltage dependent calcium channel beta 4 subunit/CACNB4 | 601949 | ****** | 2q22-q23 |
| | | | calcium channel, voltage-dependent, alpha 2/delta subunit/CACNA2D1 | 114204 | Z28613 | 7q21-q22 |
| | | | calcium channel, voltage-dependent, gamma subunit/CACNG | 114209 | NM_000727 | 17q24 |
| | | | neuronal voltage dependent calcium channel gamma subunit/CACNG2 | 602911 | NM_006078 | ****** |
| | | | ATPase, Ca++ transporting, plasma membrane 1/ATP2B1 | 108731 | NM_001682 | 12q21-q23 |
| | | | ATPase, Ca++ transporting, plasma membrane 2/ATP2B2 | 108733 | NM_001683 | 3p26-p25 |
| | | | ATPase, Ca++ transporting, plasma membrane 3/ATP2B3 | 300014 | AF060497 | Xq28 |

| | | | | | |
|---|---|---|---|---|---|
| | | | ATPase, Ca++ transporting, plasma membrane 4/ATP2B4 | 108732 | NM_001684 | 1q25-q32 |
| | | | skeletal muscle ryanodine receptor gene/RYR1 | 180901 | AH006668 | 19q13.1 |
| | | Sodium Channels | sodium channel alpha-subunit/SCN4A | 170500 | U24693 | 7q23.1-q25.3 |
| | | | type II voltage dependent sodium channel alpha 1 subunit/SCN2A1 | 182390 | M94055 | 2q23-q24.3 |
| | | | type IX voltage dependent sodium channel alpha subunit/SCN9A | 603415 | NM_002977 | 2q24 |
| | | | type I voltage dependent sodium channel alpha subunit/SCN1A | 182389 | S71446 | 2q24 |
| | | | type III voltage dependent sodium channel alpha subunit/SCN3A | 182391 | S69887 | 2q24 |
| | | | type II voltage dependent sodium channel beta subunit/SCN2B | 601327 | NM_004588 | 11q22-qter |
| | | | type VI voltage dependent sodium channel alpha subunit/SCN6A | 182392 | M55662 | 2q21-q23 |
| | | | type IV voltage dependent sodium channel alpha subunit/SCN4A | 603967 | NM_000334 | 7q23.1-q25.3 |
| | | | type V voltage dependent sodium channel alpha subunit/SCN5A | 600163 | NM_000335 | 3p24-p21 |
| | | | type VIII voltage dependent sodium channel alpha subunit/SCN8A | 600702 | ****** | 12q13 |
| | | | type II voltage dependent sodium channel alpha 2 subunit/SCN2A2 | 601219 | M55662 | 2q23-q24 |
| | | | type I voltage dependent sodium channel beta subunit/SCN1B | 600235 | NM_001037 | 19q13.1 |
| | Channels | | voltage independent neuronal sodium channel 1/ACCN1 | 601784 | NM_001094 | 17q11.2-q12 |

| | | | | |
|---|---|---|---|---|
| | voltage independent neuronal sodium channel 2/ACCN2 | 602866 | NM_001095 | 12q12 |
| Potassium Channels | cyclic nucleotide gated hyperpolarization activated potassium | 602780 | AF064876 | ****** |
| | cyclic nucleotide gated hyperpolarization activated potassium | 602781 | AF064877 | ****** |
| | voltage dependent potassium channel, KQT-like subfamily, member | 602235 | NM_000218 | 20q13.3 |
| | voltage dependent potassium channel, KQT-like subfamily, member | 602232 | AF033347 | 8q24 |
| | voltage dependent potassium channel, subfamily F, member 1/KCNF1 | 603787 | NM_002236 | 2p25 |
| | voltage dependent potassium channel, subfamily H, member 1/KCNH2 | 603305 | NM_002238 | 1q32-q41 |
| | inwardly rectifying potassium channel, subfamily J, member 4/KCNJ4 | 600504 | NM_004981 | 22q13.1 |
| | inwardly rectifying potassium channel, subfamily J, member 14/KCNJ14 | 603953 | ****** | ****** |
| | inwardly rectifying potassium channel, subfamily J, member 2/KCNJ3/HHIRK1 | 600681 | NM_000891 | ****** |
| | inwardly rectifying potassium channel, subfamily J, member 10/KCNJ10 | 602208 | ****** | 1q |
| | potassium channel, subfamily J, member 13/KCNJ13 | 603208 | AJ007557 | 2q37 |
| | voltage dependent potassium channel, subfamily K, member 1/KCNK1 | 601745 | NM_002245 | 1q42-q43 |
| | voltage dependent potassium channel, subfamily K, member 2/KCNK2 | 603219 | ****** | 1q41 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | voltage dependent potassium channel, subfamily K, member 3/KCNK3 | 603220 | NM_002246 | 2p23 |
| | | | G protein coupled potassium channel, subfamily J, member | 601534 | NM_002239 | 2q24.1 |
| | | | G protein coupled potassium channel inward rectifier/GIRK3 | 600932 | ****** | 1q21-q23 |
| | | | voltage dependent potassium channel, subfamily S, member 1/KCNS1 | 602905 | ****** | ****** |
| | | | voltage dependent potassium channel, subfamily S, member 2/KCNS2 | 602906 | ****** | 8q22 |
| | | | voltage dependent potassium channel, subfamily S, member 3/KCNK3 | 603888 | AF043472 | 2p24 |
| | | | large conductance Ca-activated K channel M type subunit 1B | 603951 | NM_004137 | 5q34 |
| | | | large conductance Ca-activated K channel M type subunit 1A | 600150 | U09384 | Chr.10 |
| | | **Chloride** | chloride channel, calcium activated, family member 1/CLCA1 | 603906 | NM_001285 | 1p31-p22 |
| | | | chloride channel, calcium activated, family member 2/CLCA2 | 604003 | NM_006536 | ****** |
| | | | chloride channel 1 , skeletal muscle/CLCN1 | 118425 | NM_000083 | 7q35 |
| | | | chloride channel 2/CLCN2 | 600570 | NM_004366 | 3q26-qter |
| | | | chloride channel 3/CLCN3 | 600580 | NM_001829 | 4q33 |
| | | | chloride channel 4/CLCN4 | 302910 | NM_001830 | Xp22.3 |
| | | | chloride channel 5/CLCN5 | 300008 | NM_000084 | Xp11.22 |
| | | | chloride channel 6/CLCN6 | 602726 | NM_001286 | 1p36 |
| | | | phospholipase A2 group | 172411 | NM_000300 | 1p35 |
| | | | phospholipase A2 group IB/PLA2G1B | 172410 | NM_000928 | 12q23-q24.1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Prostaglandins | Biosynthesis | phospholipase A2 group X/PLA2G10 | 603603 | ****** | 16p13.1-p12 |
| | | | phospholipase A2 group IVA/PLA2G4A | 600522 | U08374 | 1q25 |
| | | | phospholipase A2 group VI/PLA2G6 | 603604 | AF064594 | 22q13.1 |
| | | | phospholipase A2 group IVC/PLA2G4C | 603602 | ****** | chr. 19 |
| | | | phospholipase A2 group V/PLA2G5 | 601192 | NM_000929 | 1p36-p34 |
| | | | phospholipase C beta 3 | 600230 | U26425 | 11q13 |
| | | | lysosomal acid lipase | 278000 | NM_000235 | 10q24-q25 |
| | | | prostaglandin endoperoxide synthetase 1/COX1/PTGS1 | 176805 | AH001520 | 9q32-q33.3 |
| | | | prostaglandin endoperoxide synthetase 2/COX2/PTGS2 | 600262 | NM_000963 | 1q25.2-q25.3 |
| | | | thromboxane A synthase 1/TBXAS1 | 274180 | EG_D34613 | 7q34 |
| | | | prostaglandin D2 synthase | 602598 | M61900 | ****** |
| | | | prostaglandin I2 synthase/prostacyclin synthase/PTGIS | 601699 | EG_D83393 | 20q13 |
| | | Receptors | prostaglandin E receptor 1, EP1 subtype/PTGER1 | 176802 | NM_000955 | 19p13.1 |
| | | | prostaglandin E receptor 2, EP2 subtype/PTGER2 | 176804 | ****** | 5p13.1 |
| | | | prostaglandin E receptor 3, EP3 subtype/PTGER3 | 176806 | NM_000957 | 1p31.2 |
| | | | prostaglandin E receptor 4, EP4 subtype/PTGER4 | 601586 | NM_000958 | 5p13.1 |
| | | | prostaglandin F receptor/PTGFR | 600563 | L24470 | 1p31.1 |
| | | | prostaglandin F2 receptor negative regulator/PTGFRN | 601204 | U26664 | 1p13.1-q21.3 |
| | | | prostaglandin I2 receptor/PTGIR/prostacyclin receptor | 600022 | SEG_HUMIP | 19q13.3 |



WO 00/50639

PCT/US00/01392

| | | | | | | |
|---|---|---|---|---|---|---|
| **Inflammation** *(additional genes in Immunology)* | | **Catabolism** | 15-hydroxyprostaglandin dehydrogenase/HPGD | 601688 | NM_000860 | 4q34-q35 |
| | | | aldo-keto reductase family 1, member C2/AKR1C2 | 600450 | NM_001353 | 10p15-p14 |
| | **Platelet Activating Factor** | **Biosynthesis** | CDP-choline:alkylacetylglycerol cholinephosphotransferase | ****** | ****** | ****** |
| | | | phospholipase A2 group | 172411 | NM_000300 | 1p35 |
| | | | phospholipase A2 group IB/PLA2G1B | 172410 | NM_000928 | 12q23-q24.1 |
| | | | phospholipase A2 group X/PLA2G10 | 603603 | ****** | 16p13.1-p12 |
| | | | phospholipase A2 group IVA/PLA2G4A | 600522 | U08374 | 1q25 |
| | | | phospholipase A2 group VI/PLA2G6 | 603604 | AF064594 | 22q13.1 |
| | | | phospholipase A2 group IVC/PLA2G4C | 603602 | ****** | chr. 19 |
| | | | phospholipase A2 group V/PLA2G5 | 601192 | NM_000929 | 1p36-p34 |
| | | **Receptors** | platelet activating factor receptor/PTAFR | 173393 | M88177 | 1p35-p34.3 |
| | | **Catabolism** | platelet activating factor acetylhydrolase 1/PAFAH1 | 601690 | NM_005084 | 6p21.2-p12 |
| | | | platelet activating factor acetylhydrolase, isoform 1B, alpha | 601545 | NM_000430 | 17p13.3 |
| | | | platelet activating factor acetylhydrolase, isoform 1B, beta | 602508 | NM_002572 | 11q23 |
| | | | platelet activating factor acetylhydrolase, isoform 1B, gamma | 603074 | NM_002573 | 19q13.1 |
| | | | platelet activating factor acetylhydrolase 2/PAFAH2 | 602344 | NM_000437 | ****** |
| | | | interferon alpha1 (IFNa1) | 147660 | X02956 | 9p22 |
| | | | interferon alpha2 (IFNa2) | 147562 | ****** | 9p22 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Cytokines | Interferon | interferon beta1 (IFNb1) | 147640 | V00546 | 9p21 |
| | | | interferon beta3 (IFNb3) | 147860 | ****** | Chr.8 |
| | | | interferon omega1 (IFNw1) | 147553 | X02669 | 9p21 |
| | | | interferon gamma (IFNg) | 147570 | L07633 | 12q14 |
| | | | interferon alpha receptor 1 (IFNAR1) | 107450 | X77722 | 21q22.1 |
| | | | interferon alpha receptor 2 (IFNAR2) | 147569 | U68755 | 21q22.1-q22.2 |
| | | | interferon gamma receptor 1 | 107470 | J03143 | 6q23-q24 |
| | | | interferon gamma receptor 2 | 602376 | NM_000874 | 21q22.1 |
| | | | 2',5'-oligoadenylate synthetase | 164350 | NM_006187 | 12q24.2 |
| | | | 2',5'-oligoadenylate synthetase | 603350 | M87284 | 12q24.2 |
| | | | 2',5'-oligoadenylate synthetase | 603351 | ****** | 12q24.2 |
| | | Interleukins | interleukin 1 alpha/IL1A | 147760 | M15329 | 2q14 |
| | | | interleukin 1 beta/IL1B | 147720 | K02770 | 2q14 |
| | | | interleukin 1 receptor, type 1/IL1R1 | 147810 | M27492 | 2q12 |
| | | | interleukin 1 receptor, type 2/IL1R2 | 147811 | NM_004633 | 2q12-q22 |
| | | | interleukin 8/IL8 | 146930 | M26383 | 4q12-q13 |
| | | | interleukin 8 receptor alpha/IL8R1 | 146929 | M68932 | 2q35 |
| | | | interleukin 8 receptor beta/IL8R2 | 146928 | M94582 | 2q35 |
| | | | interleukin 10/IL10 | 124092 | M57627 | 1q31-q32 |
| | | | interleukin 10 receptor alpha/IL101 | 146933 | U00672 | 11q23.3 |
| | | | tumor necrosis factor alpha/TNFA | 191160 | X01394 | 6p21.3 |
| | | | tumor necrosis factor beta/TNFB/lymphotoxin alpha/LTA | 153440 | NM_000595 | 6p21.3 |
| | | | tumor necrosis factor receptor superfamily, member 1A/TNFRSF1A | 191190 | NM_001065 | 12p13.2 |
| | | | tumor necrosis factor receptor superfamily, member 1B/TNFRSF1B | 191191 | NM_001066 | 1p36.3-p36.2 |
| | | MIF | macrophage migration inhibitory factor (glycosylation-inhibiting | 153620 | NM_002415 | 22q11.2 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Cardiovascular** *(additional genes in Cardiovascular and* | **Clotting** | **Clotting** | factor I/fibrinogen a, alpha/FGA | 134820 | NM_000508 | 4q28 |
| | | | factor I/fibrinogen b, beta/FGB | 134830 | AH003492 | 4q28 |
| | | | factor I/fibrinogen g, gamma/FGG | 134850 | NM_000509 | 4q28 |
| | | | factor II/prothrombin | 176930 | F2 | 11p11-q12 |
| | | | factor III/thromboplastin | 134390 | NM_001993 | 1p22-p21 |
| | | | factor V/proaccelerin/labile factor | 227400 | NM_000130 | 1q23 |
| | | | factor VII/serum prothrombin conversion accelerator | 227500 | NM_000131 | 13q34 |
| | | | factor VIII/antihemophilic factor | 306700 | NM_000132 | Xq28 |
| | | | factor IX/Christmas factor/plasma thromboplastic component/hemophilia | 306900 | NM_000133 | Xq27.1-q27.2 |
| | | | factor X/Stuart factor | 227600 | NM_000504 | 13q34 |
| | | | factor XI/plasma thromboplastin antecedent | 264900 | NM_000128 | 4q35 |
| | | | factor XII/Hageman factor | 234000 | NM_000505 | 5q33-qter |
| | | | factor XIIIa1/fibrin-stabilizing factor | 134570 | NM_000129 | 6p25-p24 |
| | | | prekellikrein/Fletcher factor | 229000 | ****** | 4q35 |
| | | | kininogen/Flaujeac factor | 228960 | ****** | 3q27 |
| | | **Ion Pump** | solute carrier family 12, member 3/SLC12A3 (renal sodium/chloride transporter) | 600968 | NM_000339 | 16q13 |
| | | **Angiotensin** | renin/REN | 179820 | NM_000537 | Xq28 |
| | | | renin-binding protein/RENBP | 312420 | NM_002910 | 1q32 |
| | | | angiotensinogen/AGT | 106150 | NM_000029 | 1q42-q43 |
| | | | angiotensin II type 1 receptor/AGTR1 | 106165 | M87290 | 3q21-q25 |
| | | | guanine nucleotide binding protein (G protein), beta polypeptide 3/GNB3 | 139130 | NM_002075 | 12p13 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Renal)* | | | dipeptidyl carboxypeptidase 1 (angiotensin I converting enzyme)/ACE/DCP1 | 106180 | NM_000789 | 17q23 |
| | **Hemostasis** | **Natriuretic Peptide** | atrial natriuretic peptide precursor A/NPPA | 108780 | X01471 | 1p36.2 |
| | | | atrial natriuretic peptide precursor B/NPPB | 600295 | ****** | 1p36.2 |
| | | | atrial natriuretic peptide precursor C/NPPC | 600296 | D28874 | 2q24-qter |
| | | | natriuretic peptide receptor A/ANPRA/NPR1 | 108960 | ****** | 1q21-q22 |
| | | | natriuretic peptide receptor B/ANPRB/NPR2 | 108961 | ****** | 9p21-p12 |
| | | | natriuretic peptide receptor C/ANPRC/NPR3 | 108962 | NM_000908 | 5p14-p12 |
| | | **Endothelin** | endothelin 1/EDN1 | 131240 | NM_001955 | 6p24-p23 |
| | | | endothelin 2/EDN2 | 131241 | NM_001956 | 1p34 |
| | | | endothelin 3/EDN3 | 131242 | NM_000114 | 20q13.2-q13.3 |
| | | | endothelin converting enzyme 1/ECE1 | 600423 | NM_001397 | 1p36.1 |
| | | | endothelin A receptor isoform delta 3/EDNRA | 131243 | AF014826 | Chr.4 |
| | | | endothelin receptor type B/EDNRB | 131244 | NM_000115 | 13q22 |
| | | | cardiotrophin 1 | 600435 | ****** | ****** |
| | | | leukemia inhibitory factor/LIF | 159540 | NM_002309 | 22q12.1-q12.2 |
| | | | ciliary neurotrophic factor/CNTF | 118945 | NM_000614 | 11q12.2 |
| | | | nerve growth factor alpha | 162020 | ****** | ****** |
| | | | nerve growth factor beta | 162030 | NM_005378 | 1p13.1 |
| | | | nerve growth factor gamma sububit/NGFG | 162040 | ****** | 19q13.3 |

| | | | | | |
|---|---|---|---|---|---|
| Growth Factor | neurotrophin 3/NTF3 | 162660 | NM_002527 | 12p13 |
| | neurotrophin 5/NTF4/NTF5 | 162662 | NM_006179 | 19q13.3 |
| | neurotrophin 6 alpha/NTF6A | 604021 | NM_004149 | 19q13.3 |
| | neurotrophin 6 beta/NTF6B | 604022 | NM_004150 | chr. 19 |
| | neurotrophin 6 gamma/NTF6G | 604023 | NM_004151 | chr. 19 |
| | brain derived neurotrophic | 113505 | ****** | 11p13 |
| | growth associated protein 43/GAP43 | 162060 | NM_002045 | Chr.3 |
| | pleiotrophin/NEGF1/PTN | 162095 | AH004121 | 7q33 |
| | semaphorin 3A/SEMA3A | 603961 | ****** | ****** |
| | glial growth factor 2/neuregulin | 142445 | M94166 | 8p22-p11 |
| | neuregulin 2/NRG2 | 603818 | ****** | 5q23-q33 |
| | neurite growth promoting factor 2/NEGF2 | 162096 | ****** | 11p11.2 |
| | glial cell derived neurotrophic factor/GDNF | 600837 | NM_000514 | 5p13.1-p12 |
| | insulin-like growth factor 1/somatomedin C/IGF1 | 147440 | M11568 | 12q22-q24.1 |
| | insulin-like growth factor 2/somatomedin A/IGF2 | 147470 | NM_000612 | 11p15.5 |
| | transforming growth factor/TGFB1 | 190180 | M60315 | 9q13.1-q13.3 |
| | transforming growth factor/TGFB2 | 190220 | M19154 | 1q41 |
| | transforming growth factor/TGFB3 | 190230 | X14149 | 14q24 |
| | ciliary neurotrophic factor receptor/CNTFR | 118946 | NM_001842 | 9p13 |
| | nerve growth factor receptor/NGFR | 162010 | NM_002507 | 17q21-q22 |
| | neurotrophic tyrosine kinase receptor type 1/NTRK1 | 191315 | Y09033 | 1q21-q22 |
| | neurotrophic tyrosine kinase receptor type 2/NTRK2 | 600456 | NM_006180 | 9q22.1 |

| | | | | | |
|---|---|---|---|---|---|
| | Growth Factors and Receptors | Receptors | neurotrophic tyrosine kinase receptor type 3/NTRK3 | 191316 | NM_002530 | 15q25 |
| | | | reelin/RELN | 600514 | NM_005045 | 7q22 |
| | | | neuropilin 1/VEGF 165 | 602069 | NM_003873 | 10p12 |
| | | | neuropilin 2/VEGF 165 receptor/NP2 | 602070 | NM_003872 | 2q34 |
| | | | homolog of Drosophila | 603448 | AF071062 | 1p32-p31 |
| | | | retinoic acid receptor alpha/RARA | 180240 | NM_000964 | 17q12 |
| | | | retinoic acid receptor beta/RARB | 180220 | NM_000965 | 3p24 |
| | | | retinoic acid receptor gama/RARG | 180190 | M57707 | 12q13 |
| | | | RAR related orphan receptor A/RORA | 600825 | NM_002943 | 15q21-q22 |
| | | | RAR related orphan receptor B/RORB | 601972 | ****** | 9q22 |
| | | | RAR related orphan receptor C/RORC | 602943 | NM_005060 | 1q21 |
| | | | retinoid X receptor alpha/RXRA | 180245 | NM_002957 | 9q34.3 |
| | | | retinoid X receptor beta/RXRB | 180246 | X66424 | 6p21.3 |
| | | | retinoid X receptor gamma/RXRG | 180247 | U38480 | 1q22-q23 |
| | | | vitamin B12 receptor/cubilin/CUBN | 602997 | NM_001081 | 10p12.1 |
| | | | vitamin D receptor/VDR | 601769 | NM_000376 | 12q12-q14 |
| | | | cannabinoid receptor 1/G protein-coupled/CNR1 | 114610 | NM_001840 | 6q14-q15 |
| | | | oncogene ERBB3/HER3 | 190151 | NM_001982 | 12q13 |
| | | | homolog 3 of Drosophila Notch/NOTCH3 | 600276 | NM_000435 | 19p13.2-p13. |
| | | | SRC, FGR, YES-related oncogene FYN/FYN (receptor tyrosine kinase) | 137025 | NM_002037 | 6q21 |
| | | | receptor 1 for activated protein kinase C/RACK1 | 176981 | ****** | ****** |
| | | | insulin-like growth factor 1 receptor precursor/IGF1R | 147370 | NM_000875 | 15q25-q26 |
| | | | insulin-like growth factor 2 receptor/IGF2R | 147280 | NM_000876 | 6q26 |

| | | | | | |
|---|---|---|---|---|---|
| Growth and Differentiation (additional genes in Oncology) | | | TGF-B type I receptor/TGFBR1 | 190181 | AH006005 | 9q33-q34 |
| | | | TGF-B type II receptor/TGFBR2 | 190182 | NM_003242 | 3p22 |
| | | | TGF-B type III receptor/TGFBR3 | 600742 | L07594 | 1p33-p32 |
| | | | fibroblast growth factor receptor 1/FGFR1 | 136350 | ****** | 8p11.2-p11.1 |
| | | | fibroblast growth factor receptor 2/FGFR2 | 176943 | Y17131 | 10q26 |
| | | | fibroblast growth factor receptor 3/FGFR3 | 134934 | NM_005247 | 4p16.3 |
| | | | fibroblast growth factor receptor 4/FGFR4 | 134935 | NM_002011 | 5q35.1-qter |
| | | | laforin/EPM2A | 254780 | AF084535 | 6q24 |
| | | | steroid receptor coactivator 1/SRC1 | 602691 | NM_003743 | 2p23 |
| | | | glucocorticoid receptor interacting protein 1/GRIP1 | 601993 | NM_006540 | ****** |
| | | | nuclear receptor coactivator/AIB1 | 601937 | NM_006534 | 20q12 |
| | | | p300/CBP associated factor/PCAF | 602303 | U57317 | 3p24 |
| | | | CREB binding protein/CRB | 600140 | NM_004380 | 16p13.3 |
| | | | cyclic AMP responsive element binding protein 1/CREB1 | 123810 | NM_004379 | 2q32.3-q34 |
| | | | silencing mediator for retinoid and thyroid hormone receptors/SMRT | 600848 | NM_006312 | ****** |
| | | | retinoic and thyroid hormone receptor associated corepressor 1/TRAC1/NCOR1 | 600849 | NM_006311 | ****** |
| | | | steroid receptor RNA activator/SRA | 603819 | AF092038 | chr. 5 |
| | | | protein kinase, cAMP-dependent regulatory, type 1 beta/PRKAR1B | 176911 | ****** | 7pter-p22 |
| | | | early growth response 2/EGR2 | 129010 | NM_000399 | 0q21.1-q22.1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | neuronal growth-associated protein SCG10/SCG10 | 600621 | ****** | ****** |
| | | | apoptosis-related cysteine protease 2/caspase 2/CASP2 | 600639 | NM_006156 | 7q35 |
| | | | apoptosis-related cysteine protease 1/caspase 1/CASP1 | 147678 | L27475 | 1q22.2-q22.3 |
| | | | neuronal apoptosis inhibitory protein/NAIP | 600355 | NM_004536 | 5q12.2-q13.3 |
| | | | protein kinase C, alpha/PRKCA | 176960 | NM_002737 | 17q22-q23.2 |
| | | | protein kinase C beta-II type/PRKCB1 | 176970 | M13975 | 16p11.2 |
| | | | protein kinase C, delta/PRKCD | 176977 | NM_006254 | 3p |
| | **Signaling** | **Signaling** | protein kinase C, epsilon/PRKCE | 176975 | NM_005400 | 2p21 |
| | | | protein kinase C, gamma/PRKCG | 176980 | ****** | 19q13.4 |
| | | | protein kinase C, iota/PRKCI | 300094 | NM_002740 | Xq21.3 |
| | | | protein kinase C zeta/PRKCZ | 176982 | L14283 | ****** |
| | | | protein kinase C-like 1/PRKCL1 | 601032 | NM_002741 | 19p12 |
| | | | protein kinase C-like 2/PRKCL2 | 602549 | NM_006256 | ****** |
| | | | inositol polyphosphate-1-phosphatase/INPP1 | 147263 | NM_002194 | 2q32 |
| | | | phosphodiesterase 1A, calmodulin-dependent/PDE1A | 171890 | NM_005019 | Chr.4 |
| | | | phosphodiesterase 1B, calmodulin-dependent/PDE1B | 171891 | NM_000924 | 12q13 |
| | | | phosphodiesterase 1C, calmodulin-dependent/PDE1C | 602987 | NM_005020 | ****** |
| | | | phosphodiesterase 4A, cAMP-specific/PDE4A | 600126 | NM_006202 | 19p13.2 |
| | | | phosphodiesterase 4B, cAMP-specific/PDE4B | 600127 | NM_002600 | 1p31 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | phosphodiesterase 4C, cAMP-specific/PDE4C | 600128 | ****** | Chr.19 |
| | | | phosphodiesterase 4D, cAMP-specific/PDE4D | 600129 | NM_006203 | 5q12 |
| | | | phosphodiesterase 7A, cAMP-specific/PDE7A | 171885 | L12052 | 8q13-q22 |
| | | | phosphodiesterase 8A, cAMP-specific/PDE8A | 602972 | AF056490 | ****** |
| | | | phosphodiesterase 9A, cAMP-specific/PDE9A | 602973 | NM_002606 | 21q22.3 |
| | | | telomerase protein component 1 | 601686 | U86136 | 14q11.2 |
| | | | telomerase reverse transcriptase | 187270 | AF015950 | 5p15.33 |
| | | | telomerase RNA component | 602322 | U86046 | 3q21-q28 |
| | Oxygen Stress | Oxygen Stress | superoxide dismutase 1/SOD1 | 147450 | NM_000454 | 21q22.1 |
| | | | superoxide dismutase 2, mitochondrial/SOD2 | 147460 | X65965 | 6q25.3 |
| | | | thioredoxin-dependent peroxide reductase/TDPX1 | 600538 | NM_005809 | 13q12 |
| | | | Peptide methionine sulfoxide reductase/MSRA | 601250 | ****** | ****** |
| | | | lipoprotein, Lp(a)/LPA | 152200 | NM_005577 | 6q27 |
| | | | succinate dehydrogenase complex, subunit C, integral membrane protein/SDHC | 602413 | NM_003001 | 1q21 |
| | | | glucose-6-phosphate dehydrogenase/G6PD (mitochondrial) | 305900 | NM_000402 | Xq28 |
| | | | aldehyde oxidase 1/AOX1 | 602841 | NM_001159 | 2q33 |
| | | | mitochondrial tRNA(lys) | 590060 | M37726 | mitochondria |

| | | | | | |
|---|---|---|---|---|---|
| Mitochondrial Maintenance | Translation | mitochondrial tRNA(ser)(UCN) | 590080 | S79597 | mitochondria |
| | | mitochondrial tRNA(Gln) | 556500 | S77916 | mitochonria |
| | | mitochondrial tRNA(Thr) | 556500 | S77921 | mitochondria |
| | Protein Maturation | paraplegin (nuclear-encoded mitochondrial metalloprotease) | 602783 | NM_003119 | 16q24.3 |
| | Electron Transport | NADH-ubiquinone oxidoreductase flavoprotein 2, 24 kDa subunit/NDUFV2 | 600532 | EG_D88542 | 8p11.31-p11. |
| | | cytochrome oxidase subunit I/MTCO1 | 516030 | AF035429 | mitochonria |
| Protein Maturation and Catabolism | | cystatin B/stefin B/CSTB | 601145 | NM_000100 | 21q22.3 |
| | | cystatin C/CST3 | 105150 | NM_000099 | 20p11.2 |
| | | ubiquitin carboxy-terminal esterase L1/UCHL1 | 191342 | NM_004181 | 4p14 |
| | | ubiquitin-protein ligase E3A/UBE3A | 601623 | NM_000462 | 15q11-q13 |
| | | peptidyl-prolyl isomerase A/cyclophilin A/PPIA | 123840 | Y00052 | 7p13 |
| | | peptidyl-prolyl isomerase B/cyclophilin B/PPIB | 123841 | M60857 | chr. 15 |
| | | peptidyl-prolyl isomerase C/cyclophilin C/PPIC | 123842 | S71018 | ****** |
| | | peptidyl-prolyl isomerase D/cyclophilin D/PPID | 601753 | NM_005038 | 4q31.3 |
| | | peptidyl-prolyl isomerase E/cyclophilin E/PPIE | ****** | NM_006112. | ****** |
| | | FK506 binding protein 1A/immunophilin/FKBP1A | 186945 | NM_000801 | 20p13 |
| | | FK506 binding protein 2/immunophilin/FKBP2 | 186946 | NM_004470 | 1q13.1-q13.3 |

| | | | | |
|---|---|---|---|---|
| | | FK506 binding protein 4/immunophilin/FKBP4 | 600611 | NM_002014 | ****** |
| | | FK506 binding protein 5/immunophilin/FKBP5 | 602623 | ****** | ****** |
| | | microsomal stress 70 protein ATPase core/STCH | 601100 | U04735 | 21q11.1 |
| | | heat shock protein, DNAJ-like 1/HSJ1 | 604139 | X63368 | 2q32-q34 |
| | | heat shock protein, DNAJ-like 2/HSJ2 | 602837 | NM_001539 | ****** |
| | | alpha-B-crystallin/CRYAB | 123590 | M28638 | 1q22.3-q23.1 |
| | | heat shock transcription factor 1/HSF1 | 140580 | NM_005526 | ****** |
| | | heat shock transcription factor 2/HSF2 | 140581 | NM_004506 | ****** |
| | | heat shock transcription factor 4/HSF4 | 602438 | NM_001538 | 16q21 |
| **Cytoskeleton** | | microtubule-associated protein tau/MAPT | 157140 | NM_005910 | 17q21.1 |
| | | tubulin, alpha, brain-specific/TUBA3 | 602529 | NM_006009 | 2q |
| | | beta-tubulin gene/TUBB | 191130 | J00314 | 6p21.3 |
| | | tubulin, beta, 5/TUBB5 | 602662 | NM_006087 | ****** |
| **Adhesion** | **Adhesion** | cadherin 2/NCAD/CDH2 | 114020 | Z27440 | 18q11.2 |
| | | calpain, large polypeptide L3/CAPN3 | 114240 | NM_000070 | 5q15.1-q21.1 |
| | | neural cell adhesion molecule | 308840 | X67912 | Xq28 |
| | | neural cell adhesion molecule | 116930 | ****** | 11q23.1 |
| | | neural cell adhesion molecule | 602040 | NM_004540 | 21q21 |
| | | neural cell adhesion | 601581 | U55258 | 7q31.1-q31.2 |
| | | opioid binding cell adhesion molecule/OBCAM | 600632 | ****** | Chr.11 |
| | | nerve injury-induced protein 1/ninjurin/NINJ1 | 602062 | NM_004148 | 9q22 |
| | | protease inhibitor 12 | 602445 | NM_005025 | ****** |



WO 00/50639

PCT/US00/01392

| | | | | | |
|---|---|---|---|---|---|
| β-amyloid Metabolism | Biosynthesis | cathepsin B/b-aCTSBmyloid precursor protein secretase/CTSB | 116810 | M14221 | 8p22 |
| | | thimet oligopeptidase 1/THOP1 | 601117 | NM_003249 | 19p13.3 |
| | | amyloid beta A4 precursor | 104760 | NM_000484 | 1q21.3-q22.0 |
| | | amyloid beta A4 precursor protein-like/APPL1 | 104740 | ****** | 9q31-qter |
| | | presenilin 1/PSEN1 (membrane/adhesion) | 104311 | NM_000021 | 14q24.3 |
| | | presenilin 2/PSEN2 (membrane/adhesion) | 600759 | NM_000447 | 1q31-q42 |
| | Interacting Proteins | amyloid beta A4 precursor protein-binding, family A, member 1/APBA1/MINT1 | 602414 | NM_004664 | 9q13 |
| | | amyloid beta A4 precursor protein-binding, family A, member 1/APBA2 | 602712 | ****** | 15q |
| | | amyloid beta A4 precursor protein-binding, family B, member 1/APBB1 | 602709 | NM_001164 | 11p15 |
| | | amyloid beta A4 precursor protein-binding, family B, member 1/APBB2 | 602710 | U62325 | Chr.4 |
| | | Munc 18-1 interacting protein | 603452 | ****** | ****** |
| | | apolipoprotein E/APOE | 107741 | NM_000041 | 19q13.2 |
| | | c-jun | 165160 | J04111 | 1p32-p31 |
| | | low density lipoprotein receptor-related protein/LRP1 | 107770 | NM_002332 | 2q13.1-q13.3 |
| | | microtubule-associated protein tau/MAPT | 157140 | NM_005910 | 17q21.1 |
| | | synuclein alpha/SNCA | 163890 | NM_000345 | 4q21 |
| | | synuclein beta/SNCB | 602569 | NM_003085 | 5q35 |
| | | synuclein gamma/SNCG | 602998 | NM_003087 | 0q23.2-q23.3 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Cellular Maintainance | | | hydroxyacyl-Coenzyme A dehydrogenase, type II/HADH2 | 602057 | NM_004493 | ****** |
| | | | myeloperoxidase/MPO | 254600 | J02694 | 17q23.1 |
| | | Tau Phosphorylation State | mitogen activated protein kinase associated protein kinase 2/MAPKAPK2 | 602006 | X75346 | ****** |
| | | | protein phosphatase 2, structural/regulatory subunit A, beta/PPP2R1B | 603113 | ****** | 11q22-q24 |
| | | | protein phosphatase 2, regulatory subunit B, alpha/PPP2R5A | 601643 | NM_006243 | 1q41 |
| | | | protein phosphatase 2, regulatory subunit B, beta/PPP2R5B | 601644 | NM_006244 | 11q13 |
| | | | protein phosphatase 2, regulatory subunit B, gamma/PPP2R5C | 601645 | NM_002719 | 3p21 |
| | | | protein phosphatase 2, regulatory subunit B, delta/PPP2R1D | 601646 | NM_006245 | 6p21.1 |
| | | | protein phosphatase 2, regulatory subunit B, epsilon/PPP2R1E | 601647 | NM_006246 | 7p12-p11.2 |
| | | Catabolism | Thimet oligopeptidase | 601117 | Z50115 | 19p13.3 |
| | | | alpha-1-antichymotrypsin/AACT | 107280 | NM_001085 | 14q32.1 |
| | | | apoptosis-related cystein protease 1/caspase 1/CASP3 | 600636 | NM_004346 | 4q35 |
| | | | alpha-2-macroglobulin/A2M | 103950 | NM_000014 | 2p13.3-p12.3 |
| | | Transport | apolipoprotein A1 of HDL/APOA1 | 107680 | NM_000039 | 11q23 |
| | | | apolipoprotein A4/APOA4 | 107690 | NM_000482 | 11q23 |
| | | | apolipoprotein C1/APOC1 | 107710 | NM_001645 | 19q13.2 |
| | | | apolipoprotein D/APOD | 107740 | NM_001647 | 3q26.2-qter |
| | | | apolipoprotein E/APOE | 107741 | NM_000041 | 19q13.2 |

| | | | | | |
|---|---|---|---|---|---|
| | | | apolipoprotein J/clustrin/APOJ/CLU | 185430 | NM_001831 | 8p21-p12 |
| | | Uptake | low density lipoprotein receptor-related protein 1/LRP1 | 107770 | NM_002332 | 2q13.1-q13.3 |
| | | | low density lipoprotein receptor-related protein 2/LRP2 | 600073 | U33837 | 2q24-q31 |
| | | | low density lipoprotein receptor-related protein 5/LRP5 | 603506 | AF077820 | 11q13.4 |
| | | | low density lipoprotein receptor-related protein 8/LRP8 | 602600 | NM_004631 | 1p34 |
| | | | low density lipoprotein receptor-related protein-associated protein | 104225 | NM_002337 | 4p16.3 |
| | **Lipid Transport and Metabolism** *(additional genes below in Inflammation)* | | oxidized low density lipoprotein receptor/OLR1 | 602601 | NM_002543 | 12p13-p12 |
| | | | very low density lipoprotein receptor/VLDLR | 192977 | S73732 | 9p24 |
| | | | sortilin related receptor/SORL1 | 602005 | U60975 | 1q23.2-q24.2 |
| | | | plasma cholesterol ester transfer protein/CETP | 118470 | NM_000078 | 16q21 |
| | | | phospholipid transfer protein/PLTP | 172425 | NM_006227 | 20q12-q13.1 |
| | | Metabolism | sterol-O-acyl transferase 1/SOAT1 | 102642 | L21934 | 1q25 |
| | | | sterol-O-acyl transferase 2/SOAT2 | 601311 | ****** | chr. 12 |
| | | | HMGCoA reductase/HMGCR | 142910 | NM_000859 | 5q13.3-q14 |
| | | | pyruvate dehydrogenase complex E1-alpha subunit/PDHA1 | 312170 | L48690 | Xp22.2-p22.1 |
| | | | pyruvate dehydrogenase (lipoamide) beta subunit/PDHB | 179060 | NM_000925 | 3p13-q23 |
| | | | pyruvate dehydrogenase complex E3 subunit/DLD | 246900 | NM_000108 | 7q31-q32 |
| | | | sialyltransferase 8/GD3 | 601123 | NM_003034 | 2p12.1-p11.2 |

| | | | | | |
|---|---|---|---|---|---|
| | | | hexosaminidase A (alpha polypeptide)/HEXA | 272800 | NM_000520 | 15q23-q24 |
| | | | hexosaminidase B (beta | 268800 | M34906 | 5q13 |
| | | | lysosomal acid beta-galactosidase | 230500 | S55851 | 3p21.33 |
| | | | GM2 ganglioside activator protein/GM2A | 272750 | NM_000405 | 5q31.3-q33.1 |
| | | | UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase, polypeptide 2/B3GALT2 | 603095 | NM_003783 | 6p21.3 |
| **Myelination** | **Myelination** | | myelin proteolipid | 312080 | M27110 | Xq22 |
| | | **Receptors** | Folate Receptor Alpha/FOLR1 | 136430 | M28099 | 1q13.3-q13.5 |
| | | | Folate Receptor Beta/FOLR2 | 136425 | AF000380 | 1q13.3-q13.5 |
| | | | Folate Receptor Gamma/FOLR3 | 602469 | Z32564 | ****** |
| | | **Transporter** | Folate Transporter (SLC19A1) | 600424 | U19720 | 21q22.3 |
| | | | Vitamin B12 binding protein | 275350 | NM_000355 | 22q11.2-qter |
| | | **Glutaminati on** | folylpolyglutamate synthetase/FPGS | 136510 | M98045 | 9cen-q34 |
| | | | gamma-glutamyl hydrolase/GGH | 601509 | U55206 | ****** |
| | | | Methylenetetrahydrofolate reductase/MTHFR | 236250 | U09806 | 1p36.3 |
| | | | Dihydrofolate reductase/DHFR | 126060 | J00140 | 5q11.2-q13.2 |
| | | | 5,10-methylenetetrahydrofolate dehydrogenase, 5,10-methylenetetrahydrofolate cyclohydrolase, 10-formyltetrahydrofolate synthetase/MTHFD1 | 172460 | NM_005956 | 14q24 |
| | | | 5,10-methenyltetrahydrofolate synthetase (5-formyltetrahydrofolate cyclo-ligase)/MTHFS | 604197 | NM_006441 | Chr. 15 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | **Folate Metabolism** | **Metabolism** | phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole | 138440 | NM_000819 | 21q22.1 |
| | | | folate hydrolase 1/FOH1 | 600934 | NP_004467 | 11q14 |
| | | | 6-pyruvoyl tetrahydrobiopterin synthase/PTPS | 261640 | Q03393 | 1q22.3-q23.3 |
| | | | serine hydroxymethyltransferase 1 (soluble)/SHMT1 | 182144 | NM_004169 | 17p11.2 |
| | | | serine hydroxymethyltransferase 2 (mitochondrial)/SHMT2 | 138450 | NM_005412 | 12q13 |
| | | | Glycine aminotransferase/glycine cleavage T protein/GAT | 238310 | NM_000481 | 3p21.2-p21.1 |
| | | | 5-methyltetrahydrofolate-homocysteine methyltransferase/methionine | 156570 | NM_000254 | 1q43 |
| | | | glutamate formiminotransferase/dihydrofolate synthetase | 229100 | ****** | ****** |
| | | **Carbon Unit Activation for SAM** | methionine adenosyltransferase I, alpha/MAT1A | 250850 | NM_000429 | 10q22 |
| | | | methionine adenosyltransferase II, alpha/MAT2A | 601468 | NM_005911 | 2p11.2 |

**Table 3. ADME and Toxicology Gene List**

| Class | Pathway | Function | Name | OMIM | GID | Locus |
|---|---|---|---|---|---|---|
| | | | sucrase-isomaltase/S1 | 222900 | NM_001041 | 3q25-q26 |